8-2015

REDUCTION OF DOSE CALCULATION ERRORS FOR PATIENTS WITH METAL IMPLANTS RECEIVING PHOTON RADIATION THERAPY

Jessie Huang-Vredevoogd

Recommended Citation

Huang-Vredevoogd, Jessie, "REDUCTION OF DOSE CALCULATION ERRORS FOR PATIENTS WITH METAL IMPLANTS RECEIVING PHOTON RADIATION THERAPY" (2015). *UT GSBS Dissertations and Theses (Open Access)*. Paper 597.

REDUCTION OF DOSE CALCULATION ERRORS FOR PATIENTS WITH METAL IMPLANTS RECEIVING PHOTON RADIATION THERAPY

by

Jessie Y. Huang-Vredevoogd, B.S.

APPROVED:

____
Stephen F. Kry, Ph.D.
Advisory Professor

____
David S. Followill, Ph.D.

____
Rebecca M. Howell, PhD

____
Dragan Mirkovic, Ph.D.

____
Xinming Liu, Ph.D.

____
Francesco Stingo, Ph.D

APPROVED:

____
Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston

REDUCTION OF DOSE CALCULATION ERRORS FOR PATIENTS WITH METAL IMPLANTS RECEIVING PHOTON RADIATION THERAPY

A

DISSERTATION

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
MD Anderson Cancer Center
Graduate School of Biomedical Sciences
in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Jessie Y. Huang-Vredevoogd, B.S.

Houston, Texas

August, 2015

Acknowledgements

I would first like to thank my advisor, Dr. Stephen Kry, for all of his advice, mentorship, guidance, and support. His positive attitude, willingness to discuss ideas, and patience have been invaluable throughout this process. I would also like to thank all of my committee members, for their time and generous contributions: Dr. David Followill, Dr. Rebecca Howell, Dr. Dragan Mirkovic, Dr. Francesco Stingo, and Dr. Xinming Liu. I would like to say a special thanks to Drs. Followill and Howell for their unwavering support, encouragement, and wise career advice.

Thanks so much to the entire IROC Houston staff. I have greatly enjoyed working with such wonderful and friendly people. A special thanks goes out to the physics assistants, especially Carrie Lujano and Nadia Hernandez, for answering all my phantom questions; John Costales for his help with the many phantom machining projects; and the amazing IROC physicists, Paige Taylor, Andrea Molineu, Paola Alvarez, and Jessica Lowenstein for their generous help and support. I would also like to thank Dr. Sastry Vedam for letting me perform measurements on his machine, Dr. Peter Balter for his help with O-MAR, and Dr. Theresa Hofstede for her help with the dental phantom. To my friends and fellow students, especially Mitchell Carroll, Dana Lewis, Jackie and Austin Faught, Elizabeth McKenzie, Jessica Nute, Kiley Pulliam, and James Kerns, thanks for your friendship, many insightful discussions, and helping me avoid work when it just wasn’t happening.

Finally, the biggest thanks of all go to my parents, who have always encouraged and supported me, and of course my wonderful husband, Kevin Vredevoogd, without whom I would have surely quit ages ago.

REDUCTION OF DOSE CALCULATION ERRORS FOR PATIENTS WITH METAL IMPLANTS RECEIVING PHOTON RADIATION THERAPY

Jessie Yanshu Huang-Vredevoogd, B.S.

Supervisory Professor: Stephen F. Kry, Ph.D.

Abstract

Many patients receiving external beam radiation therapy have metal implants that can affect their treatment, and these metal implants can degrade the accuracy of dose calculations. Dose calculation errors result from limitations of modern dose calculation algorithms in modeling metal/tissue interface effects. Metals also cause streak artifacts in the computed tomography (CT) images that are used for treatment planning, and these artifacts can also degrade dose calculation accuracy. Metal based-energy deposition kernels are a potential solution for the calculation errors associated with the limitations of the convolution/superposition (C/S) dose calculation method as they better model photon interactions and scatter in metals than water-based kernels, while CT metal artifact reduction methods have the potential to decrease calculation errors associated with imaging artifacts.

In this work, several metal-based energy deposition kernels (titanium, silver, and gold) were generated and characterized. These metal-based kernels exhibited more lateral scatter, more backscatter, and less energy deposited in the forward direction than water-based kernels, implying that simply scaling water kernels according to the local density encountered is inadequate for describing photon interactions in metals. These metal kernels were then implemented into a commercial collapsed cone C/S algorithm to investigate their dosimetric impact. In comparison to water-based kernels, metal kernels resulted in better prediction of the backscatter dose enhancement upstream of metals but decreased accuracy directly downstream of metals. When used for clinical dose calculations, the dosimetric benefit of using metal kernels was small (generally limited to a 0.5% decrease in calculation error for IMRT treatment plans).

In addition to metal kernels in the dose calculation algorithm, several commercial CT metal artifact reduction methods were investigated for their success in reducing dose calculation errors: the Philips O-MAR algorithm, GE’s monochromatic gemstone spectral imaging using dual-energy CT data (GSI), and dual-energy CT imaging with a dedicated artifact reduction algorithm (MARs). Each artifact reduction method was evaluated using several implants commonly encountered in radiation oncology (hip prosthesis, spinal fixation rods, and dental fillings), and its dosimetric impact was evaluated using two clinical cases. Though not always the most successful method, O-MAR was the most consistent and thus safest candidate for all-purpose metal artifact reduction in CT simulation imaging. GSI monochromatic imaging was beneficial for smaller, low Z implants but was not able to reduce the severe artifacts caused by larger, high Z implants and had very little effect on calculation accuracy. The MARs algorithm showed great success in certain scenarios (hip prosthesis and dental fillings) but also exhibited behaviors that are undesirable (i.e., metal distortion) and can actually result in increased errors in comparison to uncorrected CT images. Consequently, the MARs algorithm should be used with abundant caution for dose calculations.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Purpose Statement

Many cancer patients receiving external beam radiation therapy have metal implants near their treatment site. Common implants encountered in radiation oncology include dental fillings, hip, leg, and arm prosthesis, spinal fixation devices, and surgical rods. In a survey conducted by the American Association of Physicists in Medicine's Task Group 63, it was found that 1-4% of radiation therapy patients have a prosthetic device that could affect their treatment.[1] Additionally, a large number of head and neck cancer patients have dental fillings that can affect their treatment.

Metal implants near the tumor can pose several challenges for treatment planning. First of all, metal objects within the patient cause imaging artifacts in the computed tomography (CT) images that are used for treatment planning and dose calculations. These streak artifacts make it difficult for the physician to visualize the diseased area and to delineate the tumor for treatment planning purposes. Furthermore, the CT images are used to assign densities to the patient's tissues for heterogeneous dose calculation. Therefore, errors in the CT numbers caused by these streak artifacts propagate to density assignment errors and subsequently dose calculation errors.

In addition to the challenges posed by CT streak artifacts, modern treatment planning systems have limited accuracy in regions near metal implants. Dose calculation algorithms are limited in their ability to accurately model the attenuation of radiation through the metal object as well as the physical processes that occur at the tissue/metal interface. Due to these challenges, the level of accuracy that can be achieved for dose calculations involving metal implants is lower than that for non-metal tissues. The purpose of this work is to investigate novel techniques to improve dose calculation accuracy for photon radiation therapy patients with metal implants.

This work has three main objectives. The first is to address dose calculation errors associated with limitations of the dose calculation algorithm by implementing novel metal-based energy deposition kernels in a commercial collapsed cone convolution/superposition dose calculation algorithm. The second is to address dose calculation errors associated with CT imaging artifacts by evaluating commercial metal artifact reduction methods. Finally, novel improvements made to the dose calculation algorithm will be combined with CT artifact mitigation methods to quantify what overall improvements in dose calculation accuracy can be achieved for patients with metal implants receiving photon radiation therapy.

1.2 Background and Significance

Types of metal implants

Common types of metal implants encountered in radiation oncology include hip prostheses, spinal fixation devices, and dental restorations. A total hip prosthesis consists of an acetabular cup, typically a polyethylene core surrounding by Co-Cr-Mo or titanium alloy outer shell, and a femoral component, which includes a femoral head portion and a stem. The femoral component is typically made of Co-Cr-Mo, a titanium alloy, or stainless steel.[1] The majority of current hip prostheses are composed of Co-Cr-Mo alloys. The diameter of the femoral head portion of the prosthesis can range from 28 mm to 54mm in diameter, with the head being hollow or solid.[2]

For patients with primary spinal tumors or spinal metastases, various types of metallic implants can be used in conjunction with surgery to stabilize the spine. These spinal stabilization implants typically include titanium pedicle screws, titanium rods, and possibly also a titanium mesh cage. The diameter of the titanium rods is typically 4-7 mm.[3, 4]

Though dental restorations are much smaller in size than spinal stabilization implants and hip prostheses, they are particularly challenging for patients receiving radiation therapy due to their high density and high atomic number, as well as the close proximity of the dental fillings to tumors and organs at risk in the head and neck regions. Dental restorations can be composed of a variety of materials, including metals such as gold (ρ=19.3 g/cc)[5], gold alloy (ρ=13.8 g/cc)[6], and various dental amalgams (ρ=8.0 g/cc).[6]

Dose perturbations caused by metal implants

The presence of a metal implant in a photon radiation therapy beam results in attenuation of the photon beam in the metal inhomogeneity as well as local interface effects. The attenuation of the photon beam in the metal affects the dose downstream of the metal at distances greater than the secondary electron range (*dmax*). The attenuation caused by the implant can be calculated accurately, given accurate knowledge about the dimensions, density, and atomic number of the metal. However, the interface effects, which occur within d_{max} of the proximal and distal interfaces, are much more challenging to model due to the effects of scattered photons from the metal implant and lack of charged particle equilibrium.[1] At the proximal or upstream tissue/metal interface, there is a dose enhancement effect caused by the metal due to backscattered photons and secondary electrons (Figure 1.1). At the distal or downstream metal/tissue interface, the dose can be increased or decreased in comparison to a homogeneous geometry depending on the photon energy. For 6MV photons, the dose downstream of the distal interface exhibits a re-buildup region, such that the dose is decreased within a distance of d_{max} of the distal interface. The magnitude of these interface effects is dependent on the density and atomic number of the metal, the physical dimensions of the metal, and the energy of the photon beam. Although modern treatment planning systems are able to accurately account of the attenuation of the metal implant, they are not able to fully model the more complex interface effects.

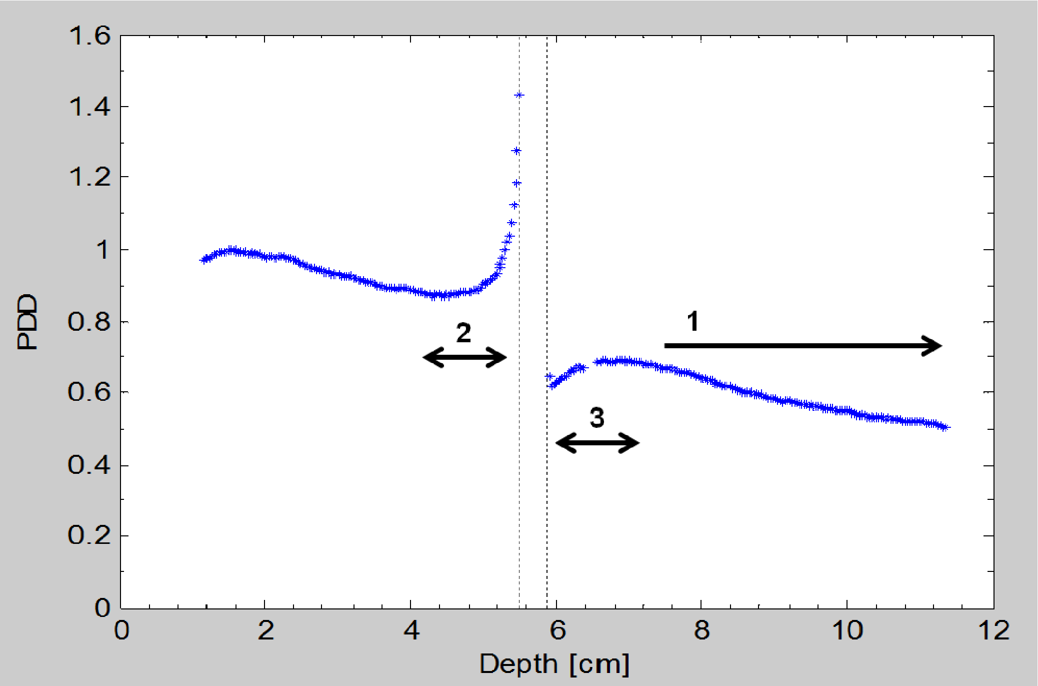


Figure 1.1: This figure illustrates the dose perturbation caused by Cerrobend metal for a 6MV photon beam. Region 1 is greater than d_{max} from the metal implant and illustrates the attenuation caused by the metal. Region 2 shows the backscatter dose enhancement effect at the proximal interface, while region 3 shows the dose perturbation at the distal interface.

Limitations of the convolution/superposition method

The concept of using kernel-based models in treatment planning systems (TPSs) was introduced and detailed by several independent investigators in the 1980's.[7-9] Since then, kernel-based or convolution/superposition (C/S) methods have become the workhorse of the modern radiation therapy clinic for photon dose calculations. Phillips Pinnacle, Nucletron Helax-TMS, Varian Eclipse, and CMS XiO are all commercial treatment planning systems with dose calculation algorithms implementing the C/S method. The C/S method separates the dose deposition process into two physically meaningful components. The energy released into the patient geometry via primary photon interactions is represented by the TERMA (total energy released per unit mass). The second component, the energy deposition kernel (EDK), describes how this energy is deposited near the interaction site by secondary electrons and scattered photons. The TERMA is convolved with the EDK in order to calculate the dose distribution.

Although the C/S method is generally able to accurately calculate the attenuation caused by metal implants in ideal cases, C/S exhibits poor accuracy near metal/tissue interfaces. C/S methods typically underestimate the backscatter enhancement at the proximately interface and overestimate the dose directly downstream of metal. For instance, Spirydovich *et al*. (2006) investigated the accuracy for a metal similar to dental amalgam and found that a commercial C/S algorithm overestimated the dose by at least 10-16% in the 3cm region downstream of the distal interface. Furthermore, the C/S algorithm was not able to model the backscatter caused by the metal and resulted in local errors as high as 30% within 5mm of the proximal interface.[10] Similarly, Wieslander and Knoos (2003) investigated the accuracy for hip prosthesis materials and found large local errors, as high as 10-25% within 3 cm downstream of the distal interface and as high as 10-15% within 1 cm upstream of the proximal interface, with the magnitude of the error depending on the metal and the photon energy.[11] It is important to note that these calculation errors occurred even in an ideal phantom geometry, in which the density of the metal was known and there were no imaging artifacts. Using a more clinically realistic geometry, Wang *et al*. 2013 investigated the accuracy of a commercial C/S algorithm for a 5 mm diameter titanium spinal rod, and found that the algorithm underestimated the backscatter dose by approximately 6%.[3] Spadea *et al* (2014), investigated dose calculation errors for a patient with gold dental fillings. In comparison to Monte Carlo results, the dose calculated using a commercial C/S algorithm resulted in a mean error of 10% in the 1cm region downstream of the filling.[12]

Limitations of CT imaging

Metal objects within the patient cause well-known imaging artifacts in the CT images used for treatment planning (Figure 1.2). CT streak artifacts are caused by two major factors: beam hardening and photon starvation. In beam hardening, the photon beam is passing through a metal object that is more attenuating than the surrounding tissue. Therefore, the portion of the X-ray beam passing through the implant is more penetrating or “harder,” meaning that the mean energy of the photon spectrum is higher because the lower energy photons are absorbed in the implant. Because of beam hardening, the attenuation measured for a particular voxel is dependent on not only the true attenuation caused by the voxel but also on the location of the voxel in a given slice, the projection angle used to measure it, and the densities of the other voxels along the projection. This dependence on location violates the underlying assumptions in the CT reconstruction method, and results in artifacts and CT number errors.[13] In the case of two dense objects, projections in which the beam passes through just one of the objects experience less beam hardening than projects that pass through both objects. This results in dark bands or streaks in between the objects.[14] Secondly, photon starvation also occurs with highly attenuating metal objects. For projections through the metal object, very few photons reach the detectors, resulting in noisy projection data and streaks in the reconstructed image. CT streak artifacts caused by metal implants degrade image quality by obscuring the relevant patient anatomy and are detrimental for both diagnostic imaging and radiation therapy treatment planning.

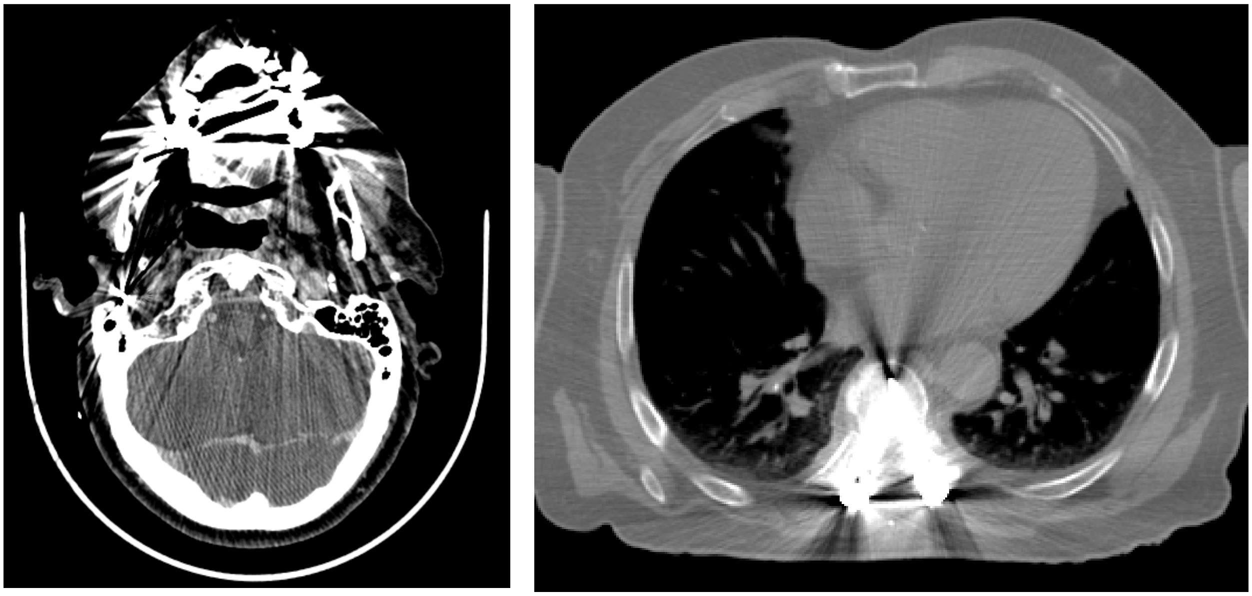


Figure 1.2: Axial CT images showing streak artifacts due to a) dental fillings and b) a spinal fixation device.

CT streak artifacts also result in inaccurate CT numbers or Hounsfield Units (HU). These HU values are used in the treatment planning system to assign an electron density for each voxel in the patient geometry in order to perform heterogeneous dose calculations. This density assignment is based on a HU-to-density calibration curve, an example of which is shown in Figure 1.3. This calibration curve is typically obtained by CT scanning a phantom containing various tissue substitute materials of known electron density. Inaccurate HU values result in inaccurate density assignment which propagates to inaccurate dose calculations.[15, 16] The magnitude of these errors depends on the severity of the streak artifacts, the area affected, and the beam arrangement for a particular treatment plan. In order to mitigate the effects of these HU errors, artifact regions can be manually contoured in the TPS so that a corrected density can be assigned to those areas (e.g., a density override of unity for tissues regions). However, this process is time consuming and requires some knowledge or assumptions of the density of the affected region.

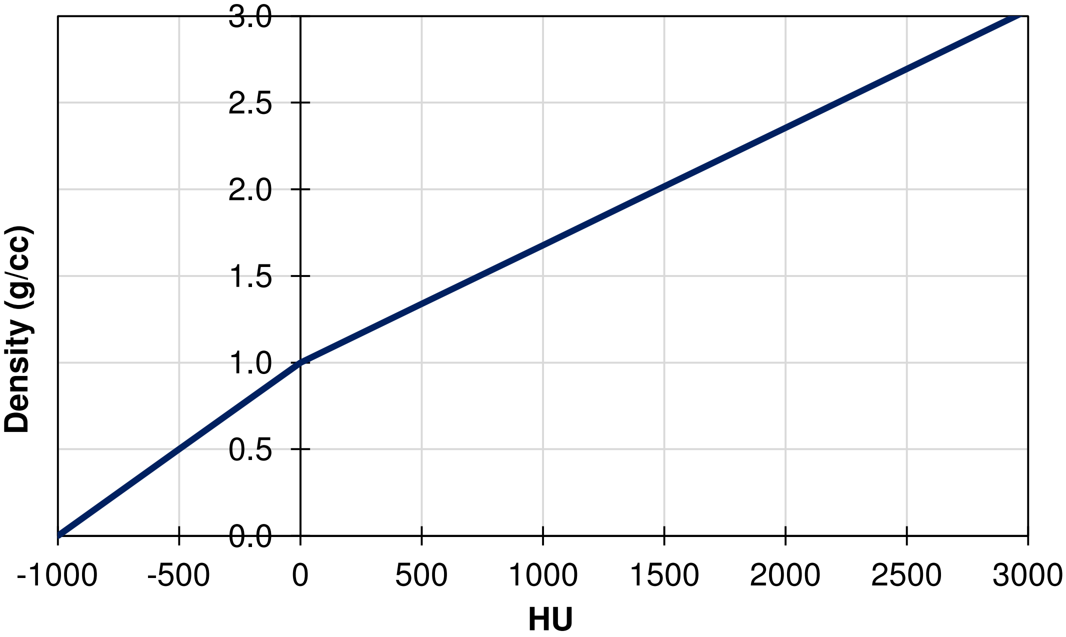


Figure 1.3: An example of a HU-to-density calibration curve. The density of a particular patient voxel is assigned based on the HU value in the CT image.

Another challenge with patients with metal implants is that the density and physical dimensions of the metal implant itself may be unknown. For cases in which the density and material composition of an implant can be obtained from the manufacturer, this information can be accounted for in the TPS, which can then accurately calculate the attenuation caused by the implant but not the local interface effects. However, when this information is unknown, the CT imaging artifacts may cause the implant to appear larger in the images than the physical dimensions of the implant. Also, it may not be possible to tell from the CT images whether a prosthetic device is hollow or solid Furthermore, the HU values reported for a high density implant will typically saturate, i.e., the maximum HU value will be assigned to the implant, which gives little to no information about the density to the implant. Extended bit-depth CT images have been investigated to overcome these issues; in these extended bit-depth images, the HU values for metal pixels do not saturate and have resulted in improved dose calculation accuracy.[17] However, extended bit-depth CT images are not yet commonplace for clinical CT imaging in radiation oncology. Without information from the CT images, estimating the density of the material requires the physicist to measure the attenuation caused by the metal, and portal images are required to see whether a prosthetic device is hollow or solid.[1] In cases in which the metal HU values are saturated and the density is not estimated via attenuation measurements, the density assignment of the implant will be heavily dependent on the CT number-to-density calibration curve, i.e., the density that corresponds to the maximum HU value on the curve.

Various artifact mitigation methods have been investigated to try to overcome the limitations of CT imaging for patients with metal implants, including iterative reconstruction methods,[18, 19] projection modification methods,[20, 21] dual energy CT,[22] and megavoltage CT. Recently, commercial metal artifact reduction options have become available for CT imaging of patients with metal implants. However, these commercial options need to be thoroughly evaluated for different types of metal implants using metrics that are relevant for radiation therapy treatment planning and dose calculation accuracy.

Clinical impact for patients with metal implants

The success of radiation therapy treatments requires a high degree of accuracy in the treatment planning and delivery process. The American Association of Physicists in Medicine's Task Group 53 suggests an overall accuracy of 5%, which includes uncertainties in the treatment planning and delivery process, as a good goal on radiobiological grounds.[23] This high level of accuracy is difficult to achieve for routine treatments, but patients with metal implants near the treatment site pose additional challenges. For patients with metal implants receiving radiation therapy, CT imaging artifacts obscure the tumor and surrounding normal tissues and result in errors in density assignment to tissues in the affected regions. Furthermore, limited information is available from the CT images about the dimensions and density of the metal implant, and the dose calculation algorithm has limited accuracy near metal/tissue interfaces. All these factors degrade the dose calculation accuracy achievable near metal implants. Dose calculation inaccuracies can lead to treatment plans that have inadequate target coverage, insufficient normal tissue sparing, and poor tumor control.

Hip prostheses affect the treatment of patients receiving pelvic irradiation, e.g., prostate patients. Because of the high dose uncertainty near metal implants, it is recommended that beam arrangements be chosen to completely or partially avoid the prosthesis when creating treatment plans for these patients. However, this beam arrangement limitation may lead to unacceptably high dose to the surrounding healthy organs. More accurate dose calculations would allow greater flexibility in treatment planning and thus better quality treatment plans for these patients.[1]

Unlike the case of hip prostheses, it is difficult to avoid beam angles that going through dental work for patients receiving head and neck radiation therapy. For arc therapies, it is especially difficult to avoid beams passing through dental materials due to the symmetric location of many head and neck tumors; restrictions on beam angles can compromise tumor coverage and increase dose to normal tissues. Furthermore, accurate dose calculation near the metal fillings is especially important for these patients because the tumor, as well as organs at risk such as the parotid glands, can be very close to the teeth and dental work. Unless the dimensions and the electron density of the high-density materials are known and accurately specified in the treatment planning system, the system will not be able to accurately model the attenuation caused by the metals, which can lead to inadequate target coverage.[5] Furthermore, the dose calculation algorithm is unable to accurately predict the dose enhancement effect due to backscatter from metal dental work, and mucositis is frequently observed adjacent to dental restorations for these patients.[6]

For patients with spinal fixation devices, stereotactic radiosurgery (SRS) may be used to treat primary spinal tumors or spinal metastases. As is the case for dental fillings, it is difficult to choose beam angles to completely avoid the spinal rods and screws. Inaccurate density assignment for the spinal hardware in the TPS can lead to underestimation of attenuation and decreased target coverage for these highly conformal treatments.[3] Furthermore, spinal SRS is a high precision treatment technique, in which the spinal cord is the dose limiting structure. It is therefore imperative that the dose delivered to the spinal cord can be accurately calculated. In cases where spinal hardware is positioned close to the spinal cord, the inability of the dose calculation algorithm to predict the backscatter effect can lead to an underestimation of spinal cord dose.

1.3 Hypothesis and Specific Aims

Accurate dose calculations are essential for achieving high quality treatment plans with adequate tumor coverage and minimal dose to nearby normal tissues. However, dose calculation accuracy is compromised for patients with metal implants due to limitations of the dose calculation algorithm and limitations of CT imaging. Thus, the aim of this study is to address these two sources of dose calculation errors near metals.

Central hypothesis: That implementation of metal kernels in the convolution/superposition dose calculation algorithm, in conjunction with metal artifact reduction methods for CT imaging, can reduce dose calculation errors near metal implants by a factor of two.

Specific Aims: This hypothesis will be tested in the following specific aims:

1. To calculate and characterize material-specific energy deposition kernels. This aim seeks to address errors caused by limitations of the convolution/superposition dose calculation algorithm.
2. To investigate and evaluate commercial metal artifact reduction methods for CT imaging. This aims seeks to address dose calculation errors caused by limitations of CT imaging.
3. To quantify gains in dose calculation accuracy due to implementation of metal kernels and metal artifact reduction methods using phantom studies. This aim is necessary to evaluate the clinical benefit of these two improvements over conventional techniques and to test the hypothesis.

In summary, the overarching goal of this study is to improve the accuracy of photon dose calculations for radiation therapy patients with metal implants. Though the challenges caused by metal implants are complex, improved dose calculation accuracy can potentially lead to more accurate treatment plans that ultimately result in better target coverage, better sparing of normal tissues, and thus better clinical outcomes for the patient.

1.4 Dissertation organization

This dissertation serves as a permanent record to document the work that was done to test the stated hypothesis. Chapters 3 through 6 are self-contained studies that each contains an introduction, methods and materials, results, discussion, and conclusion section. Chapters 3 and 4 correspond to the work done in specific aims 1 and 2 respectively, while specific aim 3 is broken up into two chapters (5 and 6). The most pertinent and interesting data is presented in each of these self-contained sections, but some supplementary figures are included in the appendices.

Chapter 2, Concept of Methods, gives detailed background information that is more in depth than the background information presented in each of the self-contained studies. Chapter 2 includes additional background information on the convolution/superposition dose calculation method and energy deposition kernels, dual energy CT, and projection modification metal artifact reduction methods.

Chapter 3 details the work done for specific aim 1, including the simulation of material-specific energy deposition kernels, their characterization, and preliminary dose calculations performed with the novel implementation of these material-specific kernels in a commercial convolution/superposition dose calculation algorithm.

Chapter 4 details the work done for specific aim 2 to evaluate three commercial metal artifact mitigation methods based on several criteria. The ability of each method to reduce CT number errors, accurately represent the size of metal implants, and reduce the severity of streak artifacts was evaluated using several anthropomorphic phantoms containing clinically realistic metal implants, including dental fillings, spinal fixation rods, and a hip prosthesis.

Chapter 5 describes the first half of the work done for specific aim 3, which is the quantification of improvements in dose calculation accuracy due to implementation of metal kernels and metal artifact mitigation methods. In this chapter, a geometric phantom geometry is used to quantify the dosimetric impact of metal kernels only, metal artifact reduction methods only, and finally metal kernels in conjunction with metal artifact reduction methods.

Chapter 6 also covers work done for specific aim 3, but instead of quantifying the dosimetric impact of metal kernels and metal artifact reduction methods in a clean, ideal phantom geometry, this chapter covers the dosimetric impact on clinical scenarios and discusses practical considerations of implementing these improvements.

Chapter 7 summarizes the research project as a whole, including major results, overall conclusions, an evaluation of the hypothesis, and a discussion of the clinical implications of this study and avenues of future work.

Chapter 2: Concept of Methods

2.1 Commercial implementations of the convolution/superposition method

In the convolution superposition (C/S) method, the energy released by photon interactions occurring in a patient is described by the TERMA (total energy released per unit mass), while the process of energy deposition via secondary electrons and scattered photons is described by the energy deposition kernel (EDK). To calculate the energy released in the patient, the incident photon spectrum is described as a function of energy, i.e., the incident fluence is categorized into several discrete energy bins. As the incident fluence travels through the patient anatomy, it is attenuated exponentially. The energy released due to these photon interactions is described by the TERMA, which also follows an essentially exponential decay with depth. The calculation of TERMA, described by Equation 2.1, is performed by using a ray-tracing technique of the incident fluence through the patient anatomy:

$$T(E) = \left(\frac{\mu}{\rho}\right) E\Phi \tag{2.1}$$

For each energy bin, TERMA is calculated as the product of the primary energy fluence ($E\Phi$) and the mass attenuation coefficient (μ/ρ).[24] For commercial systems, the mass attenuation coefficients used to calculate the TERMA and the attenuation of the energy fluence throughout the patient are typically both energy-dependent and material-dependent, i.e., different attenuation coefficients are chosen based on the density of a given voxel.[25, 26]

The TERMA for a given voxel is the convolved with the energy deposition kernel to calculate the absorbed dose for a single interaction site/voxel (convolution), and the resultant dose for all interaction sites/voxels are summed to arrive at the full dose distribution (superposition). Although the concept of the C/S method is fairly simple, its implementation in commercial treatment planning systems requires several simplifications or approximations.

Specifically, the collapsed cone approximation and the density scaling approximation for water kernels will discussed.

Collapsed cone approximation

In the collapsed cone approximation, the 4π steradian geometry surrounding a photon interaction site is discretized into several cones. All energy released from the interaction site into a particular conical volume is then transported and deposited into voxels along the cone axis.[27] This approximation is used in order to decrease dose calculation time so that the C/S method can be of practical use for treatment planning. Dose calculation algorithms that use the collapsed cone approximation in order to perform convolution calculations of TERMA and EDKs are thus referred to as collapsed cone convolution/superposition algorithms. Although this approximation is commonly used for C/S-type algorithms, different treatment planning systems differ in how they implement the collapsed cone approximation. This approximation is essentially an angular discretization of the EDK. The configuration of the cones can be based on isotropic sampling of the kernel (uniform cone size) or isoenergetic sampling of the kernel (non-uniform cone size). The number of cones also differs between different commercial algorithms. For instance, the Pinnacle CCC algorithm uses 80 cones, with 10 cones spanning the zenithal angle and 8 cones spanning the azimuthal angle,[28] while M3D's collapsed cone algorithm uses 144 isotropic cones.[25]

Energy deposition kernels

Energy deposition kernels describe the spatial distribution of energy deposition due to secondary electrons and scattered photons that are set into motion by primary photon interactions. Kernels describe the spatial distribution of energy as a function of radial distance from the primary interaction site, as well as the angular distribution of energy (e.g., how much energy is deposited in the forward direction vs. the backward direction). Kernels used in modern commercial TPSs are generated using Monte Carlo simulations in a homogeneous water geometry.[29] In these simulations, monoenergetic photons are forced to interact at the center of a spherical water geometry, and the resulting pattern of energy deposition is recorded. An example of a monoenergetic energy deposition kernel is shown in Figure 2.1. For dose calculations in commercial treatment planning systems, a single “effective mean kernel” is computed by combining monoenergetic kernels with relative weights based on the energy spectrum of the photon beam. This “effective mean kernel” is then used in the convolution calculation.[30-32]

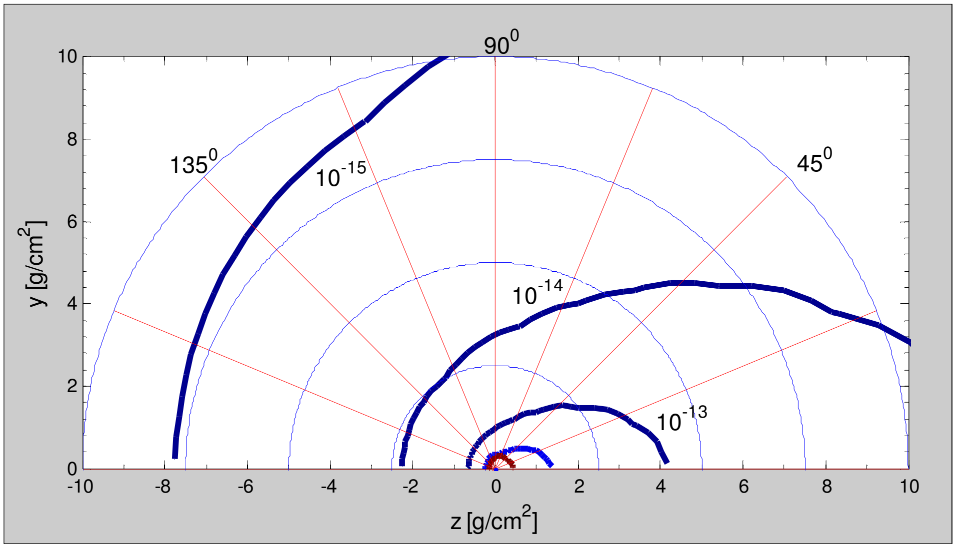


Figure 2.1: An energy deposition kernel for monoenergetic 1.0 MeV photons traveling in the +z direction (θ = 0º) direction. The incident photons interact at the origin, and the resulting distribution of energy as a function of radial distance from the interaction site and polar angle is described by the kernel.

<u>*Density scaling approximation*</u>

For dose calculations in a heterogeneous patient geometry, these water kernels are scaled (stretched or contracted in all dimensions) based on the density distribution between the interaction point and the dose deposition point.[33] Figure 2.2 illustrates the concept of density scaling of water kernels. A limitation of density scaling of water kernels is that it is only a good approximation for materials with the same atomic composition as water and thus may be inadequate to describe the physical interactions occurring in a high density, high atomic number metal implant. Woo and Cunningham (1990) investigated the validity of the density scaling method and found that it can lead to inaccuracies, by as much as 5-10% depending on the material, due to the fact that the scattering characteristics of electrons in different media are not fully taken into account.[34] Although the validity of the density scaling approximation has been tested, no studies have implemented material-specific kernels in a collapsed cone convolution/superposition method and quantified if this refinement improved dose calculation accuracy near material interfaces.

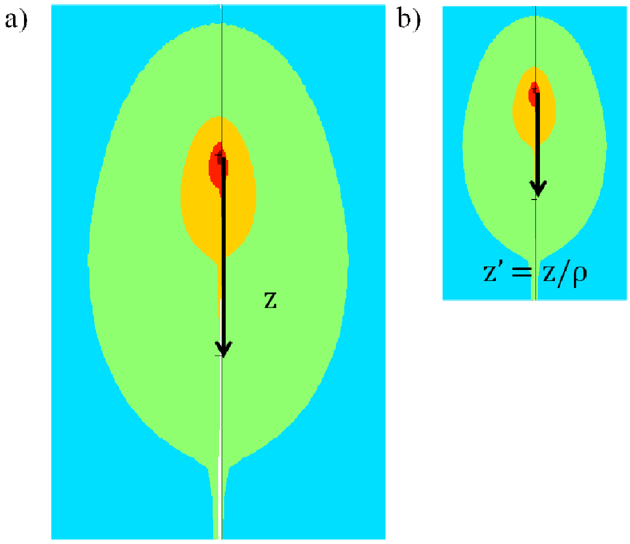


Figure 2.2: An illustration of kernel density scaling. a) shows a kernel in water (ρ = 1.0 g/cc), while b) shows a kernel in bone (ρ = 1.8 g/cc). All distances of in the kernel are scaled based on the density of the material. In this case, the high density bone causes the kernel to be contracted in all dimensions.

Implementation of metal kernels in Mobius3D

For photon dose calculations in this project, metal kernels were implemented in commercial collapsed cone convolution/superposition dose calculation algorithm (M3D, Mobius Medical Systems, LP). Although it is possible to implement various material-specific kernels to represent the heterogeneous media encountered in patient dose calculations, i.e., separate kernels for lung, tissue, bone, and metals, only two types of kernels were used in a given dose calculation in this study. For instance, for a patient with a titanium implant, titanium kernels are used for interactions occurring in the implant itself while water kernels were used for interactions occurring in the rest of the patient geometry. The implementation of metal kernels in the M3D system is based on a density threshold, i.e., for voxels in which the assigned density is greater than the threshold density (e.g., 3.0 g/cm^3), metal kernels are used to describe the energy deposition for energy released from that voxel. For voxels with density less than the threshold, water kernels are used. Kernel density scaling is still employed for voxels in which the density does not exactly match one of the two kernel materials. Both the metal and water kernels used for dose calculations were “effective mean kernels,” or polyenergetic kernels that were computed by combining monoenergetic kernels based on a single reference spectrum (e.g., 10x10 field size, on the central axis). This “effective mean kernel” method is limited in that it ignores spectral variations of the photon beam throughout the patient geometry (e.g., beam hardening with depth, off-axis softening, and beam softening for larger field sizes). In fact, not taking into account beam hardening with depth in the kernel calculation has been shown to result in errors as high as 2-5% on the central axis.[30-32]

Metal artifact reduction methods in CT imaging

<u>*Projection modification approach*</u>

Metal streak artifacts in computed tomography (CT) images are caused by a combination of beam hardening, photon starvation, scatter, and edge effects.[14, 19, 35] In the projection modification approach to metal artifact reduction (MAR), projection/sinogram data points that have been corrupted by the metal implants are identified and modified. In order to do so, the location of metal implants must be identified, either in the reconstructed images or the raw projection data. Typically, these metal data points are identified based on a threshold value and are replaced with interpolated values based on adjacent non-metal sinogram data points. Figure 2.3 illustrates the projection modification approach to metal artifact reduction. Multiple iterations may be used to correct the corrupted sinogram data such that the modified values are optimized.

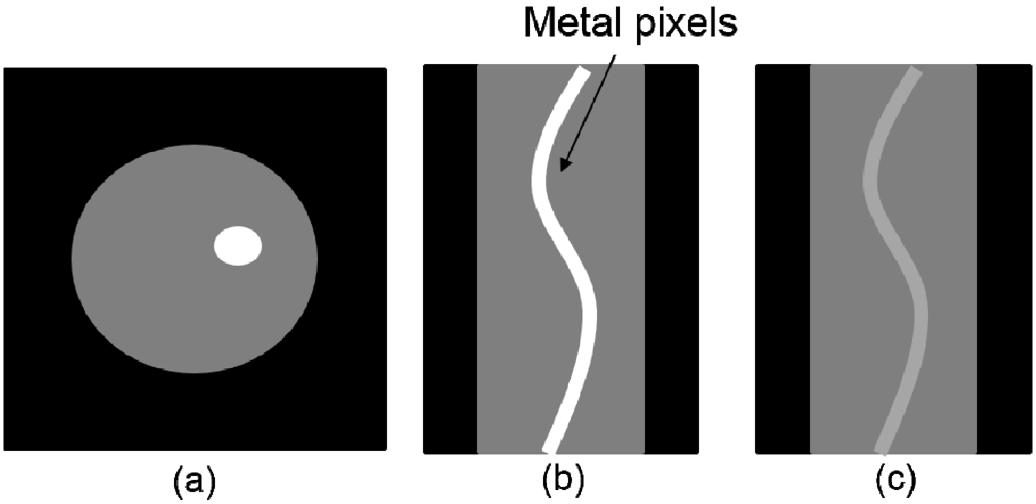


Figure 2.3: An illustration of the projection modification approach to metal artifact reduction. (b) shows the resulting sinogram data for the phantom image in (a), with the bright band of pixels being the pixels that have been corrupted by the metal implant. (c) illustrates the corrected sinogram, in which the metal pixels have been replaced with interpolated value.

O-MAR

The commercial orthopedic metal artifact reduction algorithm (O-MAR), developed by Philips Healthcare, utilizes an iterative projection modification approach. In O-MAR, the original reconstructed CT image, without any metal artifact reduction, is used as the input image for the algorithm. In the first step of the algorithm, metal pixels in the input image are identified using a threshold value, and this data is used to create a metal mask image. This input image is also segmented into a tissue pixels and non-tissue pixels. All pixels with a Hounsfield unit (HU) value near zero are classified as tissue, and their pixel values are set to the average HU value of all tissue pixels. Then, the input image, metal only image, and tissue classified image are all forward projected to obtain the corresponding sinogram data. An error sinogram is obtained from subtracting the tissue classified sinogram from the original image sinogram, and the mask image is used to remove all non-metal data points from the error sinogram. The error sinogram is then backprojected to generate the correction image, which is then subtracted from the input image. The process is repeated for multiple iterations until optimization is achieved. In the first iteration of the algorithm only, the tissue-classified image is produced not from the original uncorrected image, but from sinogram data in which the corrupted metal projections are identified, removed, and replaced with interpolated values. The O-MAR algorithm has been shown to be effective for patients with hip prostheses but can introduce secondary artifacts when the metal is located near low density tissue such as lung.[36, 37]

Dual-energy computed tomography

In dual-energy computed tomography (DECT) imaging, two datasets of the patient anatomy are acquired, one with a higher energy X-ray beam and one with a lower energy X-ray beam, e.g., 140kVp and 80kVp. From these high and low energy datasets, it is then possible to analyze energy-dependent changes in the attenuation properties of the object being imaged and to gain material-specific information. This allows the differentiation of two materials with similar linear attenuation coefficients but different material compositions. Thus, DECT has many applications in diagnostic imaging, including automatic bone removal, virtual unenhanced imaging, and characterization of urinary calcifications and abdominal lesions.[38-40]

Three types of dual-energy CT scanners are currently available commercially: a scanner with two X-ray sources, a scanner with a single X-ray source but dual detector layers, and a scanner with a single X-ray source that performs fast kilovoltage switching. The use of two X-rays sources allows better spectral separation between the high and low energy scans by adding additional filtration to each tube. However, a disadvantage of dual-source DECT is that two the datasets are acquired at slightly different times, which can limit the temporal registration of the two datasets. For DECT imaging with a single X-ray source and dual detector layers, the top layer of the detector captures the low-energy data while the bottom layer captures the high energy data. For this study, a single source dual-energy CT scanner with fast kilovoltage switching was used (Discovery CT 750HD; GE Healthcare). This dual-energy CT system rapidly switches (<0.5 ms) between 80 and 140kVp to acquire low and high energy projection data. This rapid kilovoltage switches allows good temporal registration between the high and low energy data.[22, 39]

For DECT, image reconstruction can be performed in the data domain (projection space) or the image domain. For the DECT system that performs fast kilovoltage switching, image reconstruction is performed in the data domain, which allows for greater flexibility in material decomposition. To obtain material-specific information from the low and high energy projection data, a process called "basis material decomposition" is performed. In basis material decomposition, the amount of each basis pair material (typically iodine and water) required to obtain the measured attenuation properties is calculated. With this information, it is then possible to generate two sets of images of the object, one showing the iodine density and one showing the water density of each voxel in the image (Figure 2.4). Though iodine contrast is not typically used for CT simulations in radiation therapy, iodine and water nonetheless form a good basis pair because they have sufficiently different atomic numbers, and thereby have sufficiently different attenuation characteristics as a function of energy.[41] In addition to these basis pair density images, virtual monochromatic images, that depict how an object would appear if it was imaged with a monoenergetic X-ray source, can also be generated from the basis material density images and known mass-attenuation coefficients of the basis materials.[22, 38] For the GE Discovery CT 750HD system, virtual monochromatic images can be reconstructed for energies ranging from 40 to 140keV. Along with monochromatic images, the system can also calculate effective atomic numbers for each voxel.[42]

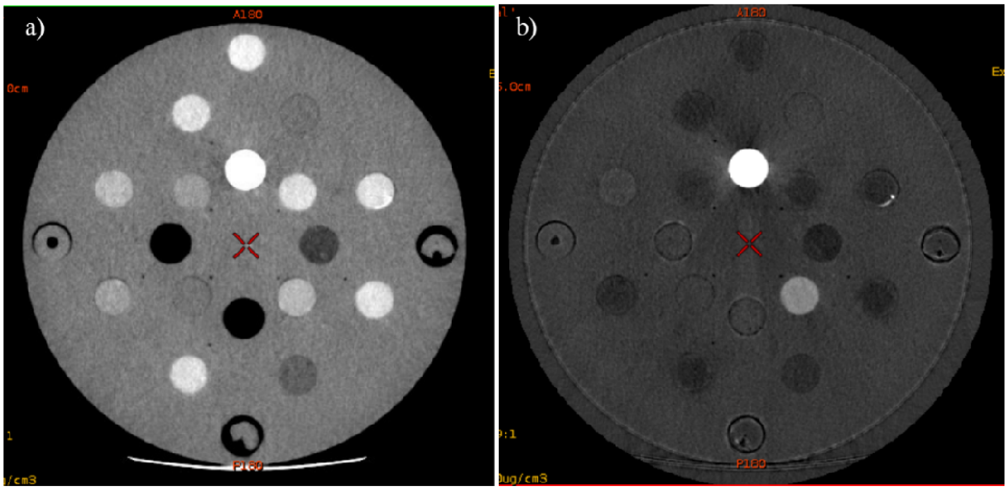


Figure 2.4: An example of basis pair density images from the GE Discovery CT 750HD DECT system. This Gammex RMI tissue characterization phantom has plugs of various tissue substitute materials, including adipose, lung, and bone. a) shows the water density map (WW = 900, WL = 1000 mg/cm^3), while b) shows the iodine density map (WW = 175, WL = 50 for units of 100 μg/cm^3).

DECT imaging with the GE Discovery CT 750HD system has been investigated as an artifact reduction method for patients with metal implants because the reconstructed virtual monochromatic images show reduced beam hardening artifacts in comparison to single-energy CT. The higher energy reconstructions especially (e.g., 140keV) have been found to improve visualization near hip prostheses.[22]

Furthermore, GE has developed a metal artifact reduction software (GSI-MARs) to be used specifically with monochromatic imaging. Conceptually, the GSI-MARs algorithm is similar to the O-MAR algorithm in that it uses projection modification for metal artifact reduction. In the MARs algorithm, metal pixels are identified in the reconstructed image based on a CT number threshold. These pixels can then be forward-projected to find the projection data points that were corrupted by the metal, and these data points can then be replaced with corrected values. However, the MARs algorithm differs from O-MAR in that the algorithm is being performed on dual-energy CT projection data and in that basis pair decomposition is then performed after the projection data have been corrected.[43] GSI-MARs therefore combines the benefits of dual-energy CT with a projection modification algorithm. GSI-MARs has shown benefits for patients with hip prostheses,[43] pedicle screws,[44] and fiducial markers.[45] However, certain pitfalls of the

GSI-MARS algorithm have also been observed. For certain metals, GSI-MARS has been show to overcorrect, resulting in the size of the metal implant appearing smaller than reality in MARs images[43] or the metal not being visible at all in the case of fiducial markers.[45]

Chapter 3: Investigation of various energy deposition kernel refinements for the convolution/superposition method

Chapter 3 was published in the journal of Medical Physics in November 2013 [J. Y. Huang, D. Eklund, N. L. Childress, R. M. Howell, D. Mirkovic, D. S. Followill and S. F. Kry, Med Phys. 40(12) 121721 (2013)]. Written permission was obtained from the journal for use of these materials in this dissertation.

3.1 Introduction

Many modern radiation therapy treatment planning systems use the convolution/superposition (C/S) method to perform photon dose calculations. The C/S method involves the convolution of the total energy released per unit mass (TERMA), calculated by attenuating the primary fluence of the photon beam through the patient (or phantom) geometry, with an energy deposition kernel (EDK).[8, 9] The EDK describes the spatial distribution of energy deposition surrounding the site of a photon interaction. Mackie *et al.*[29] generated EDKs for monoenergetic photons interacting in water by using the EGS Monte Carlo code. These kernels are currently used in the Pinnacle^3 treatment planning system (Phillips Healthcare, Andover, MA), among others.

The accuracy of the C/S method as a whole is dependent not only the accuracy of these energy deposition kernels but also on any simplifications or approximations applied to the kernels in the convolution calculation. In traditional implementations of C/S, a single, spatially invariant kernel is used for the entire dose calculation. This “effective mean kernel” is computed by combining monoenergetic kernels with relative weights based on the energy spectrum of the beam at a single location (e.g., at the surface of the patient, on the central axis). The “effective mean kernel” method ignores spectral changes of the photon beam throughout the patient geometry (e.g., beam hardening and off-axis softening) as well as spectral changes for different field sizes, and can lead to dose calculation inaccuracies.[30-32] In fact, several studies have found that higher accuracy was achieved when beam hardening of the kernels was taken into account in the convolution calculation. Dose errors caused by using non-hardened kernels were found to be as high as 2-5% on the central axis.[30-32] Although Liu *et al.* (1997) suggested that their kernel interpolation method could be used to take into account the effects of off-axis softening,[30] there have been no studies that have quantified the effects of kernel softening with larger field sizes and greater off-axis distances for a clinical photon beam.

A second simplification is density scaling of water kernels for dose calculations in heterogeneous media.[33] It has been shown that this approximation can lead to dose calculation inaccuracies for material interfaces because the lateral scattering of electrons within the inhomogeneity is not accurately modeled.[34] In fact, several studies have highlighted the limitations of the C/S method for media interfaces, especially for situations involving lateral charged particle disequilibrium in low-density media such as the lung.[27] The weakness of C/S in calculating dose near material interfaces has important clinical implications for the treatment of patients with thoracic tumors, as well as patients with high-density metals near the treatment site (e.g., dental fillings). Notably, 1-4% of radiation therapy patients have a prosthetic device that could affect their treatment.[1] C/S algorithms have been shown to overestimate the dose downstream of a metal cavity and underestimate the dose upstream of the metal due to its inability to predict the increased backscatter caused by the metal heterogeneity.[10, 11] Because of the poor dose calculation accuracy near metal prostheses, the American Association of Physicists in Medicine's Task Group 63 recommends that beam arrangements be chosen to completely or partially avoid metal prostheses even when using modern dose calculation algorithms such as the C/S algorithm. However, this beam arrangement limitation may lead to nonoptimal dose distributions, causing unacceptably high dose to organs at risk.[1] A more accurate dose calculation algorithm could allow greater flexibility in treatment planning and thus better quality treatment plans for these patients.

Although these limitations of the C/S method have been well documented, the relative importance of EDK refinements has not been explored. Determining which refinements are most important for accuracy allows for prioritization and therefore a more efficient implementation of these improvements into C/S algorithms. The purpose of this study is to generate and characterize high-resolution (i.e., those calculated with a greater number of angular and radial sample points), spatially variant, and material-specific EDKs. To this end, we performed Monte Carlo simulations to generate material-specific and high-resolution EDKs. Spatially variant polyenergetic kernels were calculated for various depths, field sizes, and off-axis distances for a clinical 6MV photon beam. We also investigated the dosimetric impact of implementing polyenergetic and material-specific kernels using simple phantom geometries.

3.2 Methods and Materials

Calculation of high-resolution kernels

The EDKnrc user code, part of the EGSnrc Monte Carlo simulation package, was used to generate high-resolution water kernels. The EDKnrc user code was derived from the EGS4 user code that was previously used to calculate water kernels by Mackie *et al*. (1988).[29, 46] For these simulations monoenergetic photons moving along the z axis were forced to interact at the center of a 60 cm radius sphere of water, and the energy deposited in each voxel was recorded. Voxels were defined by the intersection of 48 spheres (radial bins) and 144 cones (angular bins), as opposed to the 24 spheres and 48 cones used by Mackie *et al.* (1988).[29] The distance between radial boundaries was smaller near the interaction site (0.025 cm for the first few shells) and increased to 5 cm for the shells farthest from the interaction site in order to capture the higher energy deposition gradients near the origin. The angular spacing of cones was uniform (1.25°

per angular bin). Twenty-two monoenergetic photon energies were simulated between 0.1 and 40 MeV. 100 million histories were used for each simulation. Photon and electron cutoff energies (PCUT and ECUT) were set to 0.01 MeV, and the maximum fractional electron energy loss per step (ESTEPE) was 0.02.

The energy deposited in the simulation geometry was separated into primary and scatter kernels based on its origin (e.g., energy deposited due to interaction of a primary photon vs. a scattered photon). For each high-resolution kernel, the total and the primary energy fraction were calculated. The total energy fraction, F_{tot}, is defined as the total energy deposited in the spherical phantom relative to the amount of incident photon energy interacting at the origin and is obtained by summing the EDK over all voxels.[29] Similarly, the primary energy fraction, F_{prim}, is defined as the sum over all voxels of the energy deposited from charged particles set in motion from primary photons only.

Calculation of polyenergetic kernels

The Monte Carlo package BEAMnrc[47] was used to generate the primary photon spectra from a 6-MV photon beam from a 2100 Clinac linear accelerator (Varian Medical Systems, Palo Alto, CA) at various locations in a water phantom. Our BEAMnrc accelerator model consisted of the target, primary collimator, flattening filter, and jaws, and has been previously validated against measurements.[48, 49] Particle phase space data were generated at 5 different depths (d = 1.5, 5, 10, 15, and 20 cm) in a water phantom for various field sizes (5x5, 10x10, and 20x20 cm^2). Additionally, the primary photon spectra for various off-axis distances were obtained by scoring the phase space data in annular rings centered on the central axis. All phase space data were simulated with 10 million electrons incident on the target, except for simulations taken at a depth of 20 cm, which were simulated with 35 million incident electrons. The primary spectrum was isolated by excluding any photons that had interacted within the water phantom. These primary photon spectra were used to calculate the TERMA for each energy bin, which was then used as weighting factors for combining our high-resolution, monoenergetic water kernels into polyenergetic kernels that are specific to a given depth, field size, or off-axis distance.[24]

Calculation of material-specific kernels

Material-specific kernels were generated using EDKnrc, similar to the method used to generate our high-resolution water kernels, except that the simulation geometry consisted of a sphere of material rather than a sphere of water. For these simulations, the following ICRU materials[50] were used: lung, bone, titanium, silver, and gold. These materials were chosen because they can be encountered in radiation therapy (e.g., gold fillings in head and neck treatment or a titanium hip prosthesis in prostate treatment). The simulation geometry consisted of 24 spheres and 48 cones, the original kernel resolution used by Mackie *et al.* (1988).[29] However, the radii of the spheres were scaled on the basis of the density of the material such that the product of the radii and the density was the same for all materials (i.e., $r_{lung} \cdot \rho_{lung} = r_{bone} \cdot \rho_{bone}$). For each of the materials, monoenergetic kernels were calculated for various energies ranging from 100 keV to 8 MeV with 10 million histories each (ECUT = PCUT = 10 keV and ESTEPE = 0.02). Additionally, polyenergetic material-specific kernels based on the spectrum of our clinical 6-MV photon beam (10x10 cm^2 field size, depth of 1.5 cm) were calculated and compared to the corresponding polyenergetic water kernel. This energy spectrum, calculated using the methods of the previous section, was chosen because it represents a reference energy spectrum.

Investigation of "pocket of material" kernels

A further issue associated with the implementation of material-specific kernels is that 60 g/cm^2 radius spheres of heterogeneities are unrealistic in terms of patient anatomy; rather, a smaller pocket of heterogeneity surrounded by tissue is usually encountered. To address this issue, we also investigated the impact of the size of the heterogeneity on the EDK. We considered two extremes: first, as described above, we considered a uniform 60 cm g/cm^2 radius sphere of the heterogeneous material. Second, we considered spheres where the heterogeneity comprised only the central 1 mm or 5mm radius (at the interaction point), while the remainder of the sphere was comprised of water. Monoenergetic kernels were simulated with this geometry for 1.5 MeV photons for bone and titanium. These kernels will be referred to as "pocket of material" kernels.

Calculation of energy deposition kernel metrics

The effective depth of penetration $\bar{z}$, effective radial distance $\bar{r}$, and effective lateral distance $\bar{y}$ were calculated for each kernel in this study as these metrics have been previously used to characterize EDKs.[29, 51] These effective distances are a measure of the range of charged particles generated from primary photon interactions in the incident photon (*z*), the radial (*r*), and the lateral (*y*) directions. They describe how far away from the interaction site charged particles deposit energy as well as the spatial distribution of this energy. For instance, the ratio $\bar{y}/\bar{z}$ gives information about whether charged particles deposit energy more in the forward direction or the lateral direction and aids in describing the shape of the kernel. The effective distance in any given direction (*z, r,* or *y*) is essentially the expectation value of the primary EDK in that direction:

$$\bar{d} = \sum_i \sum_j \frac{d(i,j)\varepsilon_{prim}(i,j)}{F_{prim}} \tag{3.1}$$

where $\bar{d}$ indicates the effective distance to be calculated (*z*, *r*, or *y*), $d(i, j)$ is the pertinent distance from the origin to the center of the i,j^{th} voxel, and $\varepsilon_{prim}(i,j)$ is the primary EDK for the i,j^{th} voxel. To calculate $d(i,j)$, the geometric center was used to approximate the effective center of the voxel. For material-specific kernels, effective distances were also calculated using equation (3.1), with the density-scaled distance $\rho_{mat} * d$ used in place of the physical distance d.

Investigation of dosimetric impact of kernel hardening and material-specific kernels

In order to investigate the dosimetric impact of implementing spatially variant polyenergetic kernels in the C/S method, depth dose curves for our clinical 6MV photon beam were calculated using two different implementations of the collapsed cone C/S method (Mobius Medical Systems, Houston, TX):[25] one using a single polyenergetic kernel reflecting the surface spectrum ("single kernel" method) and one that represents the ideal case where all the spectral changes of the photon beam are fully taken into account. For the ideal case, a full convolution calculation using a monoenergetic kernel is performed for each energy bin ("components" method).[30, 32] Depth doses were calculated using various field sizes (5x5, 10x10, and 20x20 cm^2) using a 30x30x30 cm^3 water phantom.

In order to investigate the dosimetric impact of implementing material-specific kernels, a water phantom with a 4x4x4 cm^3 titanium cavity, representing a simplified hip prosthesis geometry, was used. The depth dose for this heterogeneous geometry was calculated using both a traditional C/S implementation that performs density scaling of water kernels and a novel implementation using titanium kernels for photon interactions occurring in the titanium cavity. The depth dose was calculated for a field size of 10x10 cm^2 for our clinical 6MV photon beam. The same calculation was also performed for a bone cavity phantom, with the same phantom geometry as the titanium phantom, using our bone kernels to describe photon interactions occurring in the bone cavity.

3.3 Results

Calculation of high-resolution kernels

In general, there was very good agreement between our high-resolution kernels and those published by Mackie *et al* (1988). This can be seen in Figure 3.1 which shows the radial dose fall-off for 1MeV and 10MeV monoenergetic water kernels.[3] The agreement between our high-resolution water kernels and the Mackie kernels was also confirmed by comparing various kernel metrics, including the total energy fraction F_{tot}, primary energy fraction F_{prim}, and effective distances $\bar{z}$, $\bar{r}$, and $\bar{y}$ (Appendix A). Although the agreement was generally good, there were some differences for the lower photon energies (<500keV), most likely attributable to improvements in the electron transport in the EGSnrc code compared to the EGS code used by Mackie *et al.*[3] Mainegra-Hing *et al.* (2005)[52] found differences in the two codes due to the different forcing algorithm and inclusion of Raleigh scattering in EGSnrc, leading to differences in the EDKs near the interaction site, especially for lower photon energies (<500 keV).

Calculation of polyenergetic kernels

The various primary beam spectra generated from Monte Carlo simulations of a Varian Clinac 2100 accelerator are shown in Figure 3.2. Notably, the mean energy of the beam spectrum was most dependent on depth. However, there were noticeable spectral changes due to field size (i.e., the spectrum is softer for larger field sizes) and off-axis distance (i.e., the spectrum is softer for larger distances from the central axis). The mean energy of each simulated primary photon spectrum is shown in Table 3.1, along with the effective distances ($\bar{z}$, $\bar{r}$, and $\bar{y}$) for each of the polyenergetic kernels. The data for spectral changes due to field size and off-axis distance shown in Table 3.1 are all for a depth of 1.5 cm. These two factors were also investigated at a depth of 20 cm, but the spectral differences due to field size and off-axis distance were less noticeable due to beam hardening at this depth. Although the spectral changes due to depth, field size, and off-axis distance are well understood, it is noteworthy that these spectral changes have caused noticeable changes in the effective distances used to characterize the polyenergetic kernels, indicating that the pattern of energy deposition is affected by these factors for a clinical photon beam.

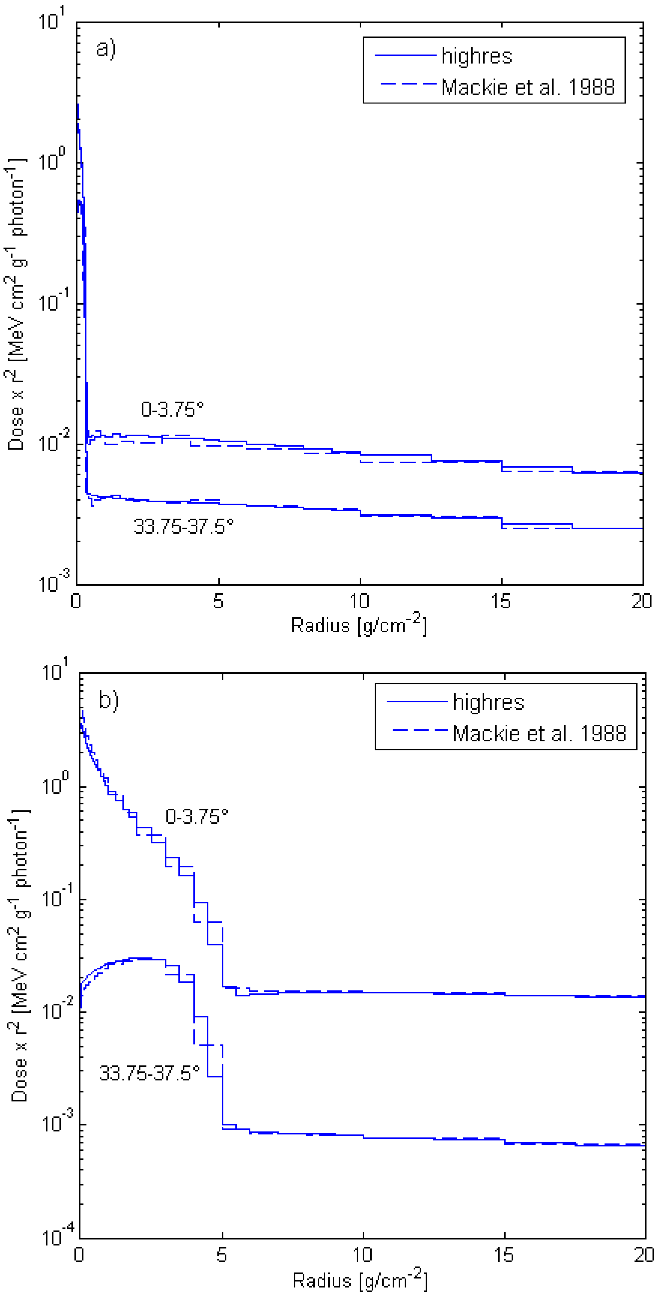


Figure 3.1: Comparison of our high-resolution energy deposition kernels for water and Mackie et al.'s kernels for (a) 1 MeV and (b) 10 MeV monoenergetic incident photons. Shown is the radial dose distribution averaged over selected angular intervals.

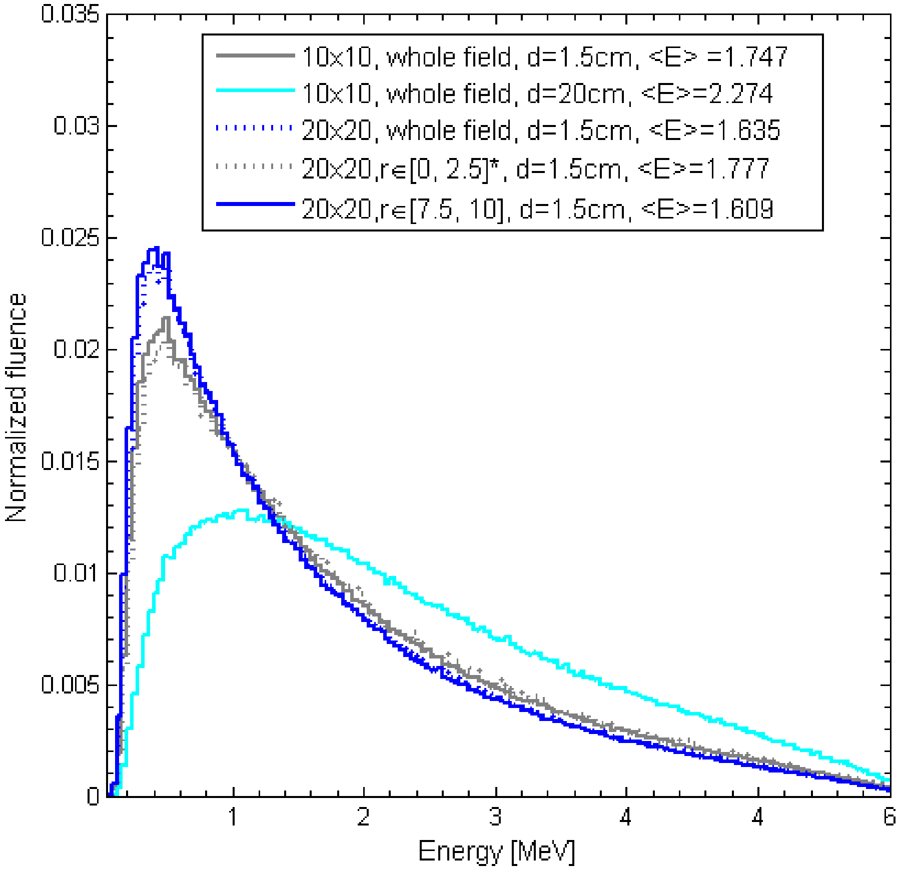


Figure 3.2: Normalized primary photon spectra for a clinical 6-MV photon beam taken at various locations in a water phantom. The mean energy of the spectrum <*E*> is given in MeV. * Indicates that the beam spectrum was taken over an annular region with an inner radius of r = 0 cm and an outer radius of r = 2.5 cm.

Table 3.1 The effective depth of penetration $\bar{z}$, effective radial distance $\bar{r}$, and effective lateral distance $\bar{y}$ for polyenergetic energy deposition kernels calculated using the energy spectrum of the primary photon beam at various locations in a water phantom. The mean energy of the spectrum is also listed. In parenthesis is the % difference with respect to the polyenergetic kernel at d = 1.5 cm for a 10x10 cm^2 field (i.e., the reference spectrum).

Variable studied	Depth (cm)	Field size ($cmxcm^2$)	Portion of field	$\bar{z}$ (cm)	$\bar{r}$ (cm)	$\bar{y}$ (cm)	Mean energy (MeV)
Depth	1.5	10x10	whole	0.327 (0.0%)	0.394 (0.0%)	0.220 (0.0%)	1.747
	5	10x10	whole	0.342 (4.5%)	0.411 (4.3%)	0.228 (3.9%)	1.853
	10	10x10	whole	0.361 (10.4%)	0.433 (9.9%)	0.239 (9.0%)	2.002
	15	10x10	whole	0.379 (15.9%)	0.454 (15.2%)	0.250 (13.7%)	2.140
	20	10x10	whole	0.397 (21.2%)	0.474 (20.2%)	0.259 (18.2%)	2.274
Field size	1.5	5x5	whole	0.335 (2.3%)	0.403 (2.2%)	0.224 (2.0%)	1.801
	1.5	10x10	whole	0.327 (0.0%)	0.394 (0.0%)	0.220 (0.0%)	1.747
	1.5	20x20	whole	0.312 (-4.7%)	0.376 (-4.6%)	0.210 (-4.1%)	1.635
Off-axis distance	1.5	20x20	r[0, 2.5]*	0.331 (1.2%)	0.399 (1.1%)	0.222 (1.0%)	1.777
	1.5	20x20	r[2.5, 5]	0.325 (-0.6%)	0.392 (-0.6%)	0.218 (-0.5%)	1.736
	1.5	20x20	r[5, 7.5]	0.317 (-3.2%)	0.382 (-3.1%)	0.214 (-2.7%)	1.673
	1.5	20x20	r[7.5, 10]	0.308 (-6.0%)	0.372 (-5.7%)	0.208 (-5.1%)	1.609

* Indicates that the beam spectrum was calculated over an annular region with an inner radius of r = 0 cm and an outer radius of r = 2.5 cm.

Calculation of material-specific kernels

The density-weighted effective distances ($\rho_{mat} \cdot \bar{z}$, $\rho_{mat} \cdot \bar{r}$, and $\rho_{mat} \cdot \bar{y}$) calculated for the material-specific kernels are listed in Table 3.2. As this table shows, the materials in this study have noticeably different values for the density-weighted effective distances. Notably, the density-weighted effective lateral distance $\rho_{mat} \cdot \bar{y}$ is different for different materials, and does not simply increase for materials of increasing density. The polyenergetic titanium kernel had the largest effective lateral distance (44% higher than the corresponding value for water). Furthermore, the ratio of $\bar{z}/\bar{y}$, which gives an indication of whether charged particles deposit more energy in the forward or the lateral direction, decreased by a factor of 3 in going from the lung to the gold polyenergetic kernel. This indicates that the shape of the kernels is changing as the density and the effective atomic number of the material is increased, with more energy being deposited in the lateral direction as opposed to the forward direction. However, the lung kernel agrees well with the water kernel, indicating that lung has similar photon interaction and scattering properties as water. Figure 3.3 illustrates the differences between various 1.5 MeV monoenergetic, material-specific kernels. In comparison to water, it can be seen that the bone and titanium kernels deposit more dose in the backward (θ = 180°) and lateral (θ = 90°) directions and less dose in the forward direction (θ = 0°).

Table 3.2: The density-scaled effective depth of penetration $\rho_{mat} \cdot \bar{z}$, effective radial distance $\rho_{mat} \cdot \bar{r}$, and effective lateral distance $\rho_{mat} \cdot \bar{y}$ for material-specific energy deposition kernels for 300 keV monoenergetic photons and a 6-MV polyenergetic beam spectrum (d=1.5cm, 10x10 field size). In parenthesis is the % difference with respect to the relevant water kernel.

Incident energy	Material	Density (g/cm^3)	$\rho_{mat} \cdot \bar{z}$ (g/cm^2)	$\rho_{mat} \cdot \bar{r}$ (g/cm^2)	$\rho_{mat} \cdot \bar{y}$ (g/cm^2)	$\bar{z}/\bar{y}$
300 keV	Water	1.00	0.016 (0.0%)	0.025 (0.0%)	0.019 (0.0%)	0.858
	Lung	0.26	0.016 (0.0%)	0.025 (0.0%)	0.019 (0.0%)	0.858
	Bone	1.85	0.015 (-7.4%)	0.026 (1.6%)	0.021 (7.9%)	0.737
	Ti	4.54	0.013 (-20.9%)	0.027 (8.8%)	0.024 (26.3%)	0.538
	Ag	10.50	0.010 (-37.4%)	0.029 (16.3%)	0.027 (43.7%)	0.374
	Au	19.32	0.007 (-58.9%)	0.026 (2.4%)	0.025 (30.5%)	0.270
6 MV spectrum	Water	1.00	0.327 (0.0%)	0.394 (0.0%)	0.220 (0.0%)	1.491
	Lung	0.26	0.328 (0.3%)	0.395 (0.1%)	0.219 (-0.3%)	1.499
	Bone	1.85	0.327 (-0.1%)	0.417 (5.8%)	0.259 (17.9%)	1.263
	Ti	4.54	0.288 (-12.1%)	0.427 (8.3%)	0.315 (43.7%)	0.912
	Ag	10.50	0.190 (-42.1%)	0.350 (-11.2%)	0.294 (33.9%)	0.645
	Au	19.32	0.117 (-64.3%)	0.265 (-32.9%)	0.237 (8.1%)	0.493

Investigation of "pocket of material" kernels

Figure 3.4 shows the radial dose falloff as a function of distance from the interaction site for our pure water, pure material, and "pocket of material" kernels for 1.5 MeV photons. We found that the dose distribution for our "pocket of material" kernels was generally very similar to that of the pure material kernels. Although there were small differences between the pure material kernels and the 1mm radius "pocket of material kernels," these differences were negligible when the size of the pocket was increased to a 5mm radius. Notably, the "pocket of material" kernels always more closely resembled the pure material kernels than the water kernels, even for very small (1mm radius) pockets of material.

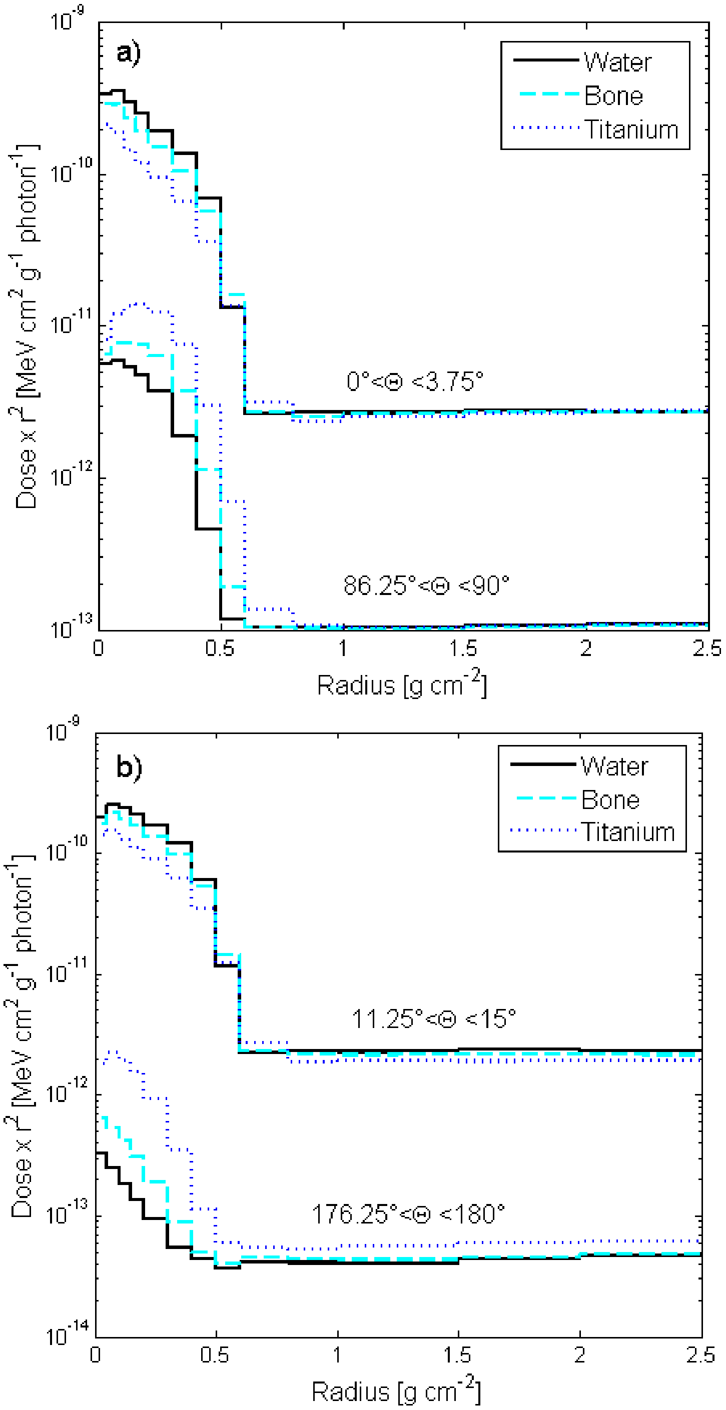


Figure 3.3: (a) and (b) Comparison of water, bone, and titanium monoenergetic kernels for 1.5 MeV incident photons. Shown is the radial dose distribution at selected angular intervals, where θ = 0° is the forward direction, θ = 90° is the lateral direction, and θ = 180° is the backward direction.

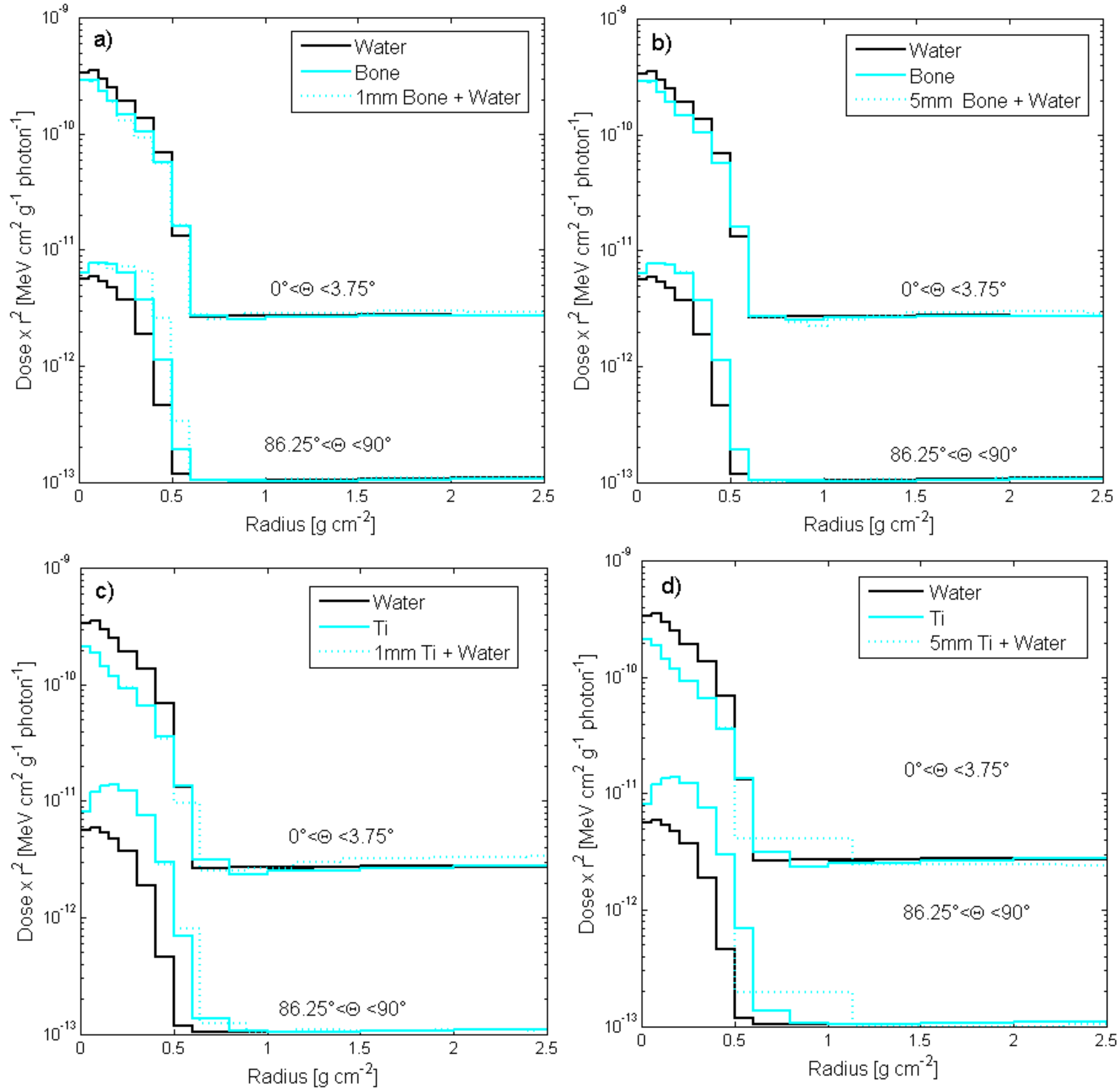


Figure 3.4: Comparison of pure water, pure material, and "pocket of material" kernels for 1.5 MeV incident photons. (a) and (b) compare the radial dose distributions for 1mm and 5mm radius pockets of bone inside water, and (c) and (d) compare 1mm and 5 mm radius pockets of titanium inside water in the forward (θ = 0°) and lateral (θ = 90°) directions.

Investigation of dosimetric impact of kernel hardening and material-specific kernels

Figure 3.5 shows percent depth dose (PDD) curves for a water phantom calculated using two different C/S implementations: one using a single polyenergetic kernel ("single kernel" method) and one that fully takes into account spectral changes in the photon beam ("components" method). The "components" method resulted in a harder PDD curve (i.e., larger PDD value at a given depth) than the "single kernel" method. The PDD value was 2.1% to 5.8% larger at a depth of 25 cm and 1.0 to 2.5% larger at a depth of 10 cm for the "components" method, with the discrepancy being larger for smaller field sizes.

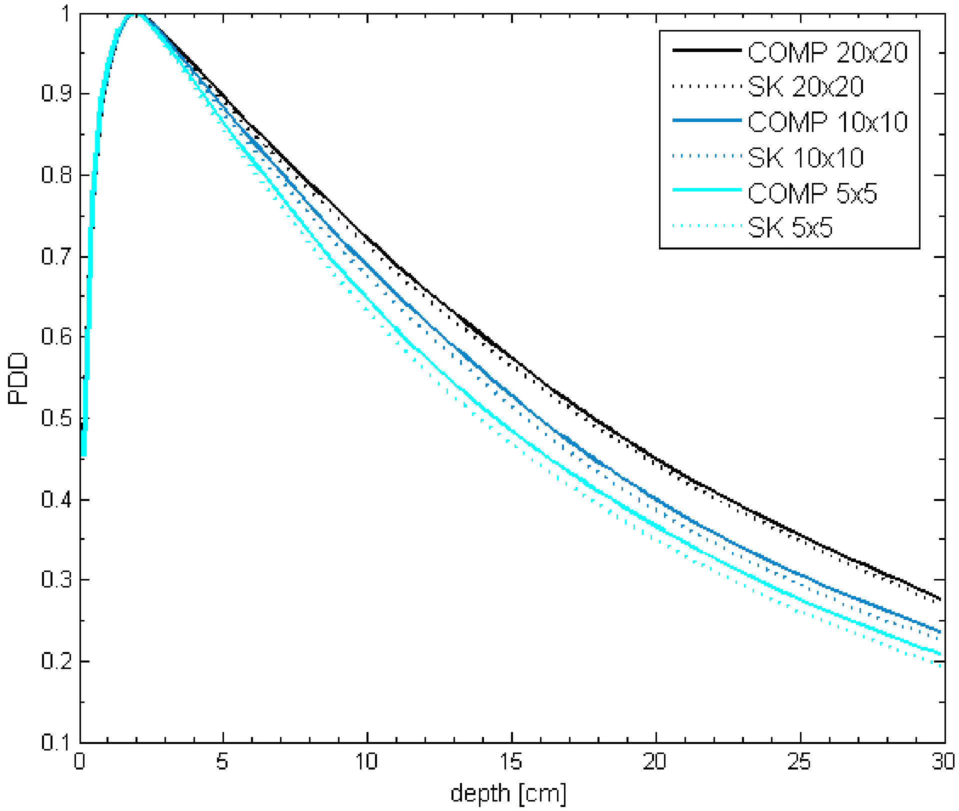


Figure 3.5: Percent depth dose curves for a water phantom calculated using a single polyenergetic kernel ("SK") and the components method ("COMP") for several field sizes.

Figure 3.6 shows PDD curves for the titanium cavity phantom and the bone cavity phantom calculated using both traditional density scaling of a water kernel, as well as a novel implementation using material-specific kernels. For the titanium cavity phantom, implementation of material-specific kernels resulted in 4.9% higher dose upstream of the metal (i.e., higher backscatter dose) and 8.2% lower dose downstream of the metal. Similar trends were found for the bone cavity phantom but the dose differences were of smaller magnitude. Materials-specific kernels resulted in 1.0% higher dose upstream and 1.5% lower dose downstream of the bone cavity.

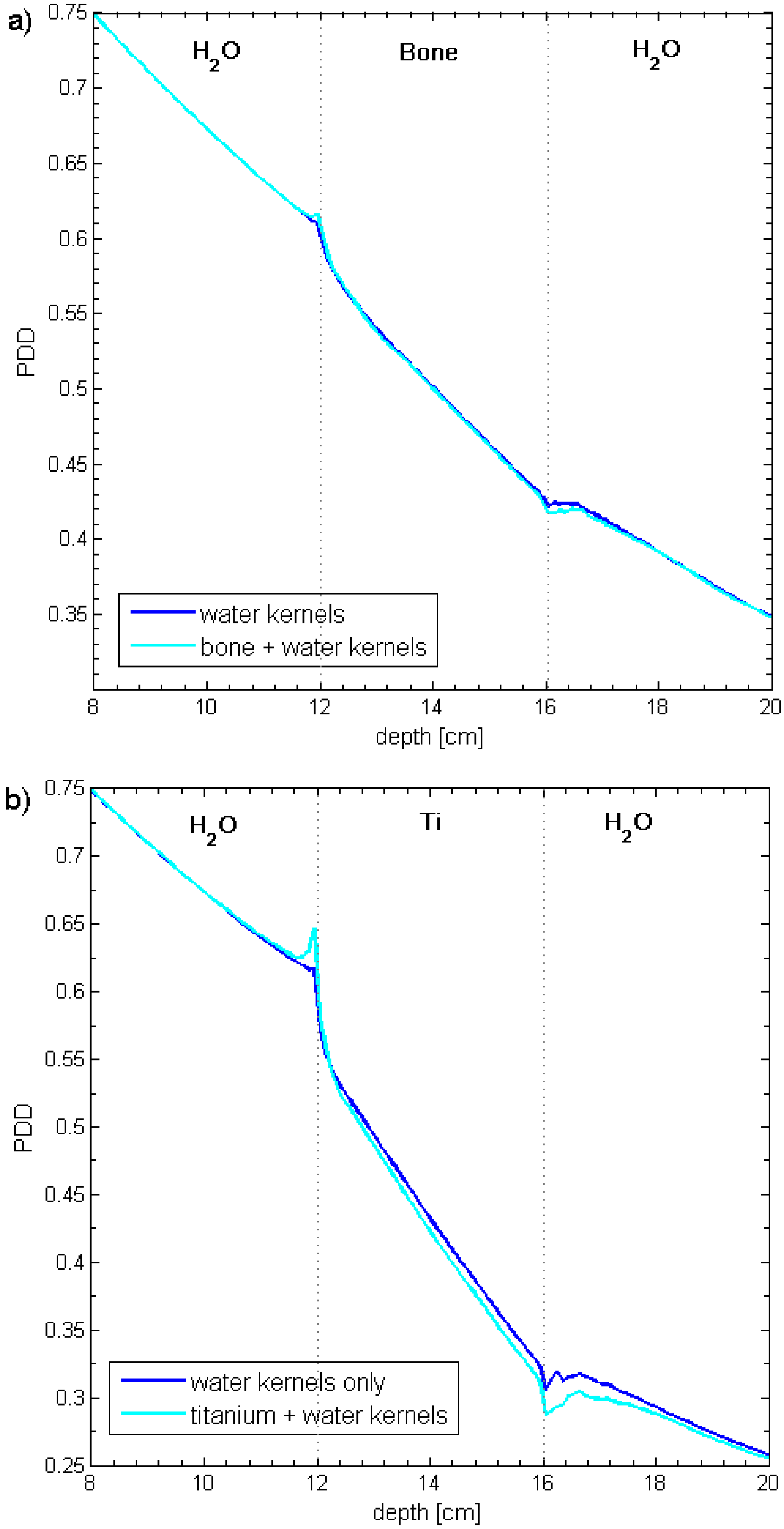


Figure 3.6: Percent depth dose curves for the (a) bone cavity phantom and (b) titanium cavity phantom (heterogeneity extends from z=12 cm to z=16 cm) calculated using traditional density scaling of water

kernels (“water kernels”) and novel implementation of material-specific kernels (“bone + water kernels” and “titanium + water kernels”).

3.4 Discussion

In this study, we generated high-resolution, polyenergetic, and material-specific photon EDKs. Our high-resolution kernels agreed well with the original Mackie *et al.* kernels.[29] However, we did observe differences near the interaction site for the lower photons energies (<500 keV), most likely attributable to improvements in electron transport in the EGSnrc code compared with the EGS code.[53] Because the kernels were generally smoothly varying functions, interpolation of the kernels is likely a good approximation. Therefore, we believe that incorporating higher-resolution kernels into the C/S method will not appreciably improve the calculation accuracy, with the possible exception of near material interfaces and in the buildup region.

The impact of the photon spectrum used to calculate polyenergetic kernels was also investigated. Specifically, the effects of beam hardening with depth and beam softening with greater field size and off-axis distance were investigated. Among these three factors, beam hardening with depth was the most dominant. The effective depth of penetration of charged particles $\bar{z}$ for the polyenergetic kernels increased by 21%, going from a depth of 1.5 cm to a depth of 20 cm for a 10x10 cm^2 field size (Table 3.1). Melcalfe *et al.* (1990)[24] also found that the effective range of charged particles increased for depth-hardened kernels in comparison to the surface kernel. Although the use of depth-dependent polyenergetic kernels has already been studied and shown to improve dose accuracy, no studies, to our knowledge, have investigated the effects of beam softening for larger field sizes and greater off-axis distances. Our data showed that the effective depth of penetration of charged particles $\bar{z}$ can change as much as 7% for polyenergetic kernels calculated for various field sizes and off-axis distances (Table 3.1). Although these spectral differences are smaller than those attributed to beam hardening with depth, they are still appreciable and taking them into account can improve dose calculation accuracy. We also found that the spectral differences due to field size and off-axis distance are smaller at greater depths due to beam hardening across the entire field. Therefore, kernel corrections for these factors may only be needed at shallower depths. We also compared the depth dose in a water phantom calculated with a single polyenergetic kernel against that calculated using the components method, which fully takes into account spectral changes in the photon beam. We found that the error associated with not performing any type of kernel hardening correction is approximately 2-6% for PDD values at a depth of 25 cm. These values are consistent with dose errors reported in previous studies.[30, 32] The disadvantage of the components method is the increased computation time. The components method increased the computation time by a factor of 8 in this study in comparison to the single kernel method. Liu et al. (1997)[30] proposed a kernel hardening correction method, using interpolation to calculate the mean polyenergetic kernel as a function of depth, and found that this method was able to match the accuracy of the components method while only increasing the computation time by 50%. This kernel interpolation method may be a practical method of accounting for kernel hardening. However, because only spectral changes as a function of depth are taken into account, further studies need to be performed to look at the accuracy of this method off-axis.

It is also important to note that although the dose differences due to kernel hardening found in this study were fairly small, implementation of a kernel hardening correction can still have an impact on accuracy. For instance, in the TPS commissioning process, implementation of kernel hardening can potentially lead to beam model parameters that more closely reflect reality and thus a more accurate overall beam model. While our future work includes quantifying how the implementation of depth-hardened kernels affects various beam modeling parameters for fully-commissioned beam models, this is outside the scope of the current study.

For our material-specific kernels, we found that the effective depths of penetration of charged particles were very similar for water and lung, indicating that the density scaling approximation is appropriate for this material. However, we found that as the density and effective atomic number differed more from water, the shape of the material-specific kernel also differed more from water. For higher density, higher *Z* materials, a greater proportion of the energy is deposited in the lateral direction. For our polyenergetic, material-specific kernels the effective lateral range of charged particles in bone was 18% higher than that of water, and in titanium was 44% higher than that of water (Table 3.2). Furthermore, we found that the shape of bone and especially metal kernels was poorly described by density scaling and thus density scaling may not be appropriate for these materials. Density scaling cannot accurately predict the lateral scatter and angular energy deposition caused by high-*Z*, high density materials. Woo and Cunningham[34] also investigated the validity of the density scaling method and found it was too simple to adequately describe the scatter of electrons in different media and thus caused dose inaccuracies in regions greatly affected by inhomogeneities. The inadequacy of the density scaling method for metals is a possible reason why the C/S method has been shown to be inaccurate near metal/tissue interfaces, overestimating the dose downstream and underestimating the backscatter dose upstream of metals. In this study, we found that the implementation of a metal-specific kernel resulted in increased backscatter dose and decreased dose downstream of a metal cavity, thus decreasing both of these dose errors associated with traditional C/S.

In reality, heterogeneities in patients are relatively small pockets within tissue, but these pockets are of different sizes depending on the situation. Developing distinct material kernels for pockets of different sizes would be impractical and difficult to implement. Fortunately, we found that the dose distribution from “pocket of material” kernels was very similar to that of pure material kernels. That is, implementing material-specific kernels simulated with large volumes of pure material (which are easy and consistent to calculate) are appropriate for describing pockets of material (which are realistic), and will give better accuracy than density scaling of water kernels.

Incorporating these kernel refinements into a commercial C/S algorithm increases the computation time.[30] Thus, kernel refinements must be prioritized based on their relative importance. Based on our data, the differences in the kernels due to depth hardening is comparable to the differences between water and metal kernels. Of the various factors investigated, these two have the most dominant effects on the characteristics of the kernels, suggesting that incorporation of depth-specific and metal-specific kernels will give the greatest increase in accuracy. Incorporation of field size and off-axis position dependence will have some impact, comparable to implementation of bone kernels. These parameters should therefore be considered for inclusion into C/S algorithms as secondary refinements.

An issue that may arise in the implementation of material-specific kernels is the number of distinct materials to be included and how to handle intermediate-density materials (i.e., materials for which there is no material-specific kernel). To investigate this issue, we fit our polyenergetic material-specific kernels to the analytical function described by Ahesnjo *et al.* (1989).[27] This analytical function describes the primary and scatter energy deposition components using a double exponential function with an inverse square correction for radial distance. Although our kernels were well described by this function, the fitting coefficients did not appear to have a predictable relationship with density, and thus material-specific kernels cannot be implemented by simply using an analytical function with fitting coefficients that scale with density. Practically, material-specific kernels will most likely be implemented such that the material kernel that is closest in density to the actual material found in the patient geometry is used. For instance, there would be a single metal kernel in the C/S algorithm and any material with a density above a certain threshold would use this kernel. This approximation will likely be an adequate one because, although the kernel shape does change in a complex way with respect to density, it does not change very drastically. Implementing a finite set of material-specific kernels (i.e., water, bone, and metal kernels only) for patient geometries with many types of materials is also the subject of future research.

3.5 Conclusions

The use of depth-hardened and material-specific energy deposition kernels are refinements to the convolution/superposition (C/S) method that have the potential to improve accuracy while taking advantage of the vast experience that clinicians and researchers have with this dose calculation algorithm. Density scaling of water kernels may be inadequate for describing the photon interaction characteristics of high-density, high Z metals. Implementation of a metal kernel in a collapsed cone C/S algorithm resulted in increased backscattered dose upstream of a metal cavity and lower dose downstream of the cavity, mitigating errors associated with traditional C/S. Taking into account spectral changes of the photon beam in the kernel convolution, namely beam hardening with depth, resulted in a harder depth dose curve in water. Implementation of depth-hardening kernel correction has the potential to affect beam modeling parameters obtained in the treatment planning system commissioning process and thus the overall accuracy of the system.

Chapter 4: An evaluation of three commercially available metal artifact reduction methods for CT imaging

Chapter 4 was published in the journal of Physics in Medicine and Biology in January 2015 [J. Y. Huang, J. R. Kerns, J. L. Nute, X. Liu, P. A. Balter, F. C. Stingo, D. S. Followill, D. Mirkovic, R. M. Howell, and S. F. Kry, Phys. Med. Biol. 60 1047-67 (2015)]. Written permission was obtained from the journal for use of these materials in this dissertation.

4.1 Introduction

Patients requiring computed tomography (CT) imaging routinely have metal implants, and these implants cause well-known imaging artifacts. These artifacts, when severe, not only degrade diagnostic image quality, but also complicate the radiation therapy treatment process. In radiation therapy treatment planning, CT images are used for delineating targets and critical organs, defining the treatment geometry, and assigning densities for heterogeneous dose calculations. For treatment planning, CT imaging artifacts make it difficult for the physician to confidently delineate the tumor and surrounding organs and cause errors in CT numbers (expressed in Hounsfield units [HU]), which can propagate to density assignment errors and subsequently dose calculation errors.[15, 16, 54]

CT streak artifacts are caused by a combination of beam hardening, photon starvation, scatter, edge effects, and patient motion.[14, 19, 35] Various metal artifact reduction algorithms have been investigated in an effort to overcome the various causes of metal artifacts.[12, 19-21, 55-57] Recently, commercial metal artifact reduction options have become available for CT imaging of patients with metal implants. This study focuses on three artifact mitigation methods: the algorithm for orthopedic implants (O-MAR) developed by Philips Healthcare (Cleveland, OH), monochromatic Gemstone Spectral Imaging (GSI) using dual-energy CT data without any additional metal artifact reduction post-processing applied, and GSI monochromatic imaging with metal artifact reduction software applied (MARs).

The Philips O-MAR algorithm is an iterative projection modification solution, whereby the data corrupted by streak artifacts are identified and corrected based on uncorrupted projection data. The O-MAR algorithm segments the original reconstructed image into metal and tissue pixels and uses these data to calculate a correction image.[37] O-MAR was designed primarily for orthopedic implants but has also been found to be effective for non-orthopedic metals, such as dental fillings.[36] Although O-MAR has begun to be implemented in radiation oncology clinics, few published studies have evaluated its performance for treatment planning.[58] Li *et al.* (2012) found that CT number accuracy, noise, and image quality were improved with the use of O-MAR; however, their study was limited in that it focused solely on patients with hip prostheses receiving radiation therapy for prostate cancer. Glide-Hurst *et al.* (2013) investigated the O-MAR algorithm, in conjunction with extended-bit depth, for several patient cases with various types of orthopedic implants.[17]

Unlike the Philips O-MAR algorithm, which is a software-only approach for conventional CT data, GSI monochromatic imaging is a fundamentally different approach to metal artifact reduction that uses dual-energy CT data. The HD750 Discovery system (GE Healthcare, Milwaukee, WI) uses a single X-ray source that rapidly switches between two kilovoltage settings (80 and 140 kVp) to acquire projections using alternating high and low energy X-ray spectra.[22, 59] With projection data acquired at two different energies, it is then possible to generate synthesized virtual monochromatic images that depict how an object would look if it were imaged using a monoenergetic X-ray source.[42] Although they are not truly monochromatic images, these virtual monochromatic images show reduced beam hardening artifacts in comparison to conventional polyenergetic images.[38, 40, 60] The GSI dual-energy CT system has shown promise in diagnostic imaging for patients with orthopedic prostheses,[43] spinal screws,[44] and fiducial markers.[45] Although there has been some interest in using GSI dual-energy CT for radiation therapy treatment planning,[61] no studies to the authors' knowledge have to date performed a thorough evaluation of this dual-energy CT system as an artifact mitigation method for treatment planning purposes.

To further reduce artifacts, GE has developed metal artifact reduction software (MARs) specifically for use with GSI monochromatic imaging that addresses the photon starvation aspect of metal streak artifacts.[43] It should be noted that for this study GSI virtual monochromatic imaging and GSI imaging with MARs will be evaluated separately as two different metal artifact reduction methods. Although GSI images and MARs images can be reconstructed from the same acquired projection data, it is the authors' opinion that the two methods are sufficiently different and should be considered separately. GSI virtual monochromatic imaging shows reduced beam hardening artifacts without MARs, while application of MARs is a post processing step on the monochromatic images that can further reduce artifacts but can also affect the appearance of metal implants.[43, 44]

Although studies of these artifact mitigation methods have been published, these have typically focused on only a single type of implant, such as a hip prosthesis, have evaluated only a single metal artifact mitigation method, or have evaluated only those metrics that are important for diagnostic imaging. Therefore, a more extensive evaluation of these commercial methods would be valuable, especially one that includes the various implants that are commonly encountered in radiation therapy and criteria that are relevant for treatment planning and dose calculation accuracy. This would provide users with information about the merits and pitfalls of each method. Thus, the purpose of this study was to evaluate the success of these three commercial artifact mitigation methods based on several criteria: the accuracy of CT numbers in regions of interest, accuracy in the representation of the size of metal objects, and reduction in the severity of streak artifacts. To perform this evaluation, we used four different phantoms: a geometric tissue characterization phantom and three anthropomorphic phantoms equipped with metal implants, including a hip prosthesis, dental fillings, and spinal fixation rods.

4.2 Methods and Materials

Imaging protocol

Several phantoms were used to evaluate the artifact mitigation methods. All phantoms were scanned using both the Philips Brilliance (Cleveland, OH) and the Discovery™ CT750 HD (GE Healthcare, Milwaukee, WI) scanners. Each phantom was scanned without metal to acquire an artifact-free image ("baseline scan") and with metal ("metal scan"). For the Philips scanner, each phantom was scanned using 120kVp and then reconstructed with and without the O-MAR algorithm. For the GE scanner, each phantom was scanned using polyenergetic imaging, i.e., 120kVp, and with dual-energy mode, which allows for monochromatic image reconstruction. For the GSI dual-energy scans, monochromatic images were generated at two different energies, 70keV and 140keV. 70keV was chosen because it closely matches the contrast-to-noise ratio of the conventional 120kVp scans;[62] 140keV was chosen because it is the highest energy available and thus allowed us to evaluate the GSI system at a wide range of reconstruction energies. Image sets were generated at both monochromatic energies (70keV and 140keV) with and without the MARs algorithm. CT protocols were matched between the two different vendors as closely as possible based on various acquisition and reconstruction parameters; these parameters are listed in Table 4.1 for all the phantom scans in this study. All reconstructed images were 12-bit depth images. In summary, phantom images were obtained using conventional imaging protocols ("Philips 120kVp" and "GE 120kVp") and the three artifact mitigation methods that we investigated: O-MAR, GSI monochromatic imaging ("GSI 70keV" and "GSI 140keV"), and GSI monochromatic imaging with MARs ("MARs 70keV" and "MARs 140keV").

Table 4.1: Scan protocols for all phantom scans using the Philips Brilliance and the HD750 GE Discovery CT scanners. All protocols are helical scans.

Phantom	Protocol	Pitch	mA	Tube rotation time (s)	Filter	SFOV (cm)	Slice thickness (cm)	DFOV (cm)	Recon kernel	CTDIvol (mGy)
RMI phantom	Philips 120kVp	0.938	375	0.75	B	60	3.0	36	B	19.9
	GE 120kVp	0.969	450	0.5	Medium	50	2.5	36	Standard	21.2
	GSI preset #35	0.969	630	0.5	Medium	50	2.5	36	Standard	20.2
Head phantom	Philips 120kVp	0.688	202	0.75	B	60	3.0	25	B	26.2
	GE 120kVp	0.516	280	0.6	Medium	50	2.5	25	Standard	26.0
	GSI preset #32	0.516	375	0.6	Medium	50	2.5	25	Standard	26.0
Pelvic phantom	Philips 120kVp	0.938	263	0.75	B	60	3.0	36	B	13.9
	GE 120kVp	0.984	265	0.6	Medium	50	2.5	36	Standard	12.7
	GSI preset #41	0.984	360	0.6	Medium	50	2.5	36	Standard	12.8
Spine phantom	Philips 120kVp	0.938	300	0.75	B	60	3.0	50	B	15.9
	GE 120kVp	0.969	265	0.7	Medium	50	2.5	50	Standard	17.1
	GSI preset #25	0.969	375	0.7	Medium	50	2.5	50	Standard	17.0

RMI phantom

The RMI 467 tissue characterization phantom (Gammex, Middleton, WI) contains several interchangeable plugs that mimic various heterogeneous tissues, including adipose, solid water, liver, cortical bone, and lung. To mimic a range of prostheses materials, several custom-made metal cylindrical plugs were also used with this phantom: aluminum, stainless steel, titanium, and Cerrobend. The titanium plug is actually a bundle of 7 smaller rods. Metal and tissue substitute plugs were 2.8 cm in diameter. The phantom was scanned with a single metal plug in a peripheral location (for each of the four materials) and with two metal plugs in a bilateral configuration (Figure 4.1). For the bilateral configuration, titanium and stainless steel plugs were chosen because these materials are commonly used for hip prostheses.[1] Data analysis was performed using the central slice of the RMI phantom image set, and all scans were repeated three times to investigate the reproducibility of the various imaging techniques and artifact mitigation methods. The RMI phantom images were analyzed for all three evaluation criteria: CT number accuracy, metal diameter accuracy, and severity of streak artifacts.

Anthropomorphic phantoms

Because the results of the RMI phantom were dependent on the arrangement of metals and tissue substitute plugs in the phantom, more clinically applicable results were desired. Therefore, anthropomorphic phantoms were also used to evaluate the artifact mitigation methods in a geometry that more closely represented actual clinical conditions. Three anthropomorphic phantoms were chosen to represent common metal implants encountered in radiation oncology, as described below. For these anthropomorphic phantoms, each method was evaluated on the basis of its ability to reduce the severity of streak artifacts.

Pelvic phantom with hip prosthesis. This anthropomorphic pelvic phantom was designed by the Radiological Physics Center (RPC; Houston, TX) and contains structures mimicking the prostate, bladder, and rectum contained in a centrally located water-filled imaging insert.[63] The outer portion of the phantom, which is also water-filled, contains structures mimicking the femoral heads. For this study, the phantom was modified to hold a cobalt-chromium hip prosthesis (6.84 g/cm^3). The phantom was imaged with and without the hip prosthesis.

Head phantom with dental fillings. To investigate the effectiveness of the metal artifact reduction methods on artifacts caused by dental work, a CIRS Model 606 head phantom (Computerized Imaging Reference Systems, Inc, Norfolk, VA) with articulating lower jaw, tongue, teeth, and air cavities was used. Two teeth in the lower jaw could be removed and interchanged with teeth containing dental restoration materials. In addition, a metal crown was taped on top of one of the original non-metal teeth. This phantom was scanned both with metal (with two fillings and crown) and without metal.

Thoracic phantom with spine stabilization rods. This anthropomorphic phantom was designed by the RPC and contains lung and heart structures, as well as a target structure in the left lung. A spine insert is usually included that contains structures representing the spinal cord, bone, and esophagus. For this study, the spinal insert was replaced with a high impact polystyrene insert. This insert included two titanium rods (9.5 mm diameter), mimicking spinal fixation rods. This phantom was scanned both with and without metal rods. For the baseline scan without the rods, the holes in the rectangular insert were filled in with high impact polystyrene rods.

<u>Data analysis</u>

CT number accuracy. CT number accuracy was quantified for the various artifact mitigation methods by using scans of the RMI phantom. The mean CT number of select tissue substitute plugs, chosen for their location within regions of streak artifacts, was obtained by using a 14 mm diameter region of interest (ROI) centered on the plug. The mean CT number over the ROI was measured on the baseline scan (no metal; $\overline{HU_{baseline\ image}}$) and on every scan that included metal ($\overline{HU_{metal\ image}}$), including both uncorrected and corrected metal images. The CT number error $\overline{\Delta HU}$ was then calculated for each metal scan using Eq. (4.1).

$$\overline{\Delta HU} = \overline{HU_{metal\ image}} - \overline{HU_{baseline\ image}} \tag{4.1}$$

Metal diameter accuracy. The diameters of the stainless steel plug scanned with the RMI phantom and of the titanium rods scanned with the anthropomorphic thoracic phantom were calculated by identifying the metal pixels in the CT image using a threshold HU value (half the maximum metal HU value) and calculating the metal area in the image. For both metals, the metal area was obtained from five images, and the average metal area was used to calculate the diameter of the metal plug/rod. This calculated diameter was then compared to the physical diameter of the plug/rod obtained by electronic caliper measurements.

Severity of streak artifacts. The severity of streak artifacts was quantified for all four phantoms. The analysis included the entire phantom in the image plane (excluding any regions of air and the metal implant). For the RMI phantom, the analysis was performed for a central slice in the phantom image set, whereas for the anthropomorphic phantoms, multiple slices spanning the metal implant were analyzed. To quantify the severity of streak artifacts, the baseline image was registered to the metal image using rigid, intensity-based image registration in MATLAB. After image registration, an HU error map was created by subtracting the baseline image from the metal image. All pixels with an HU error of >40 HU were considered to be "bad pixels," i.e., pixels in which the HU accuracy was negatively affected by the presence of streak artifacts. This 40 HU threshold was chosen because it corresponds to approximately a 0.03 g/cm^3 density assignment error for water. This density assignment error was found to result in approximately 1%-2% dose calculation errors for 6MV photon treatments and is the electron density tolerance level recommended by Kilby *et al.* (2002). For each phantom image set analyzed, the percentage of bad pixels (pixels with HU error > 40 HU) was calculated. In addition, to take into account the magnitude of the HU errors of the bad pixels, an error metric M_{error} was also calculated using Eq. (4.2), where $\overline{\Delta HU_{badpixels}}$ is the mean absolute CT number error of the bad pixels in the image.

$$M_{error} = \frac{\%\, bad\ pixels}{100} * \overline{\Delta HU_{badpixels}} \qquad (4.2)$$

For example, an M_{error} of 40 could correspond to 50% of the phantom pixels being bad (HU error >40), where the mean absolute CT number error of these bad pixels was 80 HU. For the anthropomorphic phantoms, for which several image planes were analyzed, average values for % bad pixels and M_{error} across the slices analyzed are reported.

4.3 Results

RMI phantom

CT number accuracy. Mean HU errors ($\overline{\Delta HU}$) due to streak artifacts for conventional CT imaging as well as the artifact mitigation methods can be seen in Figure 4.1 for select tissue substitute regions of interest and various metal configurations of the RMI phantom. The mean HU errors are grouped by imaging technique. In each case, three tissue equivalent inserts were evaluated. These inserts were selected due to their position within the area most strongly affected by the streak artifacts. A decreased absolute value of $\overline{\Delta HU}$, i.e., decreased bar height, indicates an improvement in CT number accuracy. In general, the standard deviation of $\overline{\Delta HU}$ for three repeated scans was low (<10 HU) for all of the artifact mitigation methods for all scan configurations of the RMI phantom, indicating good reproducibility of the artifact mitigation methods. This reproducibility is depicted by the tight error bars in Figure 4.1.

For Philips O-MAR, the HU accuracy was generally improved by application of the algorithm. The improvement in CT number accuracy was the most dramatic for the scan with bilateral steel and titanium plugs (Figure 4.1c). For this metal configuration, for the solid water material, the absolute HU error ($\overline{\Delta HU}$) decreased from >400 HU to 64 HU after O-MAR. However, the exception to this improvement in CT number accuracy was lung materials, which were made significantly worse with the application of O-MAR in some cases, as can be seen in both Figures 4.1a and 4.1b. For instance, $\overline{\Delta HU}$ for lung increased from 52 HU to 72 HU due to O-MAR for the stainless steel plug scan (Figure 4.1b).

For GSI imaging, monochromatic 70keV reconstructions gave similar or slightly worse values for $\overline{\Delta HU}$ as polyenergetic 120kVp imaging with the GE scanner. In contrast, monochromatic 140keV reconstructions generally gave lower $\overline{\Delta HU}$ values than 120kVp imaging, often substantially improving HU accuracy (e.g., Figure 4.1c). Application of the MARs algorithm further decreased HU errors, except for the scan with the titanium plug (Figure

4.1a), for which MARs increased HU errors for the 140keV image set. The largest increase in $\overline{\Delta HU}$ caused by MARs occurred for the LN-450 lung material, for which MARs resulted in an increase in absolute $\overline{\Delta HU}$ from 12 to 49 for 140keV imaging (Figure 4.1a). However, for the bilateral scan with titanium and stainless steel plugs, MARs was very successful in decreasing the large CT number errors for the tissue equivalent inserts medial to the two metal plugs where the artifacts were most severe. For instance, MARs resulted in a decrease in absolute $\overline{\Delta HU}$ from >300 HU to 38 HU for the solid water plug for 70keV monochromatic imaging (Figure 4.1c).

Mean CT number errors caused by the aluminum plug were small ($\overline{\Delta HU} < 20$) for all of the imaging techniques and algorithms studied, while CT number errors caused by the Cerrobend plug were very similar to those of the stainless steel plug (Appendix B).

Metal diameter accuracy. All of the imaging techniques were able to represent the diameter of the stainless steel plug/rod with an accuracy of ±1.4 mm (approximately 2 pixels). Of note, whereas all of the other imaging techniques tended to overestimate the stainless steel rod diameter, the MARs algorithm caused an underestimation of the diameter (Table 4.2).

Severity of streak artifacts. The error metric M_{error} (Eq. 4.2), the fraction of bad pixels (those whose error is >40 HU) multiplied by the mean absolute CT number error of these bad pixels, is shown in Figure 4.2 for all scan configurations of the RMI phantom. Philips O-MAR, GSI 140keV, MARs 70keV, and MARs 140keV all successfully decreased the severity of streak artifacts, with MARs 140keV imaging being the most successful method overall, based on this metric.

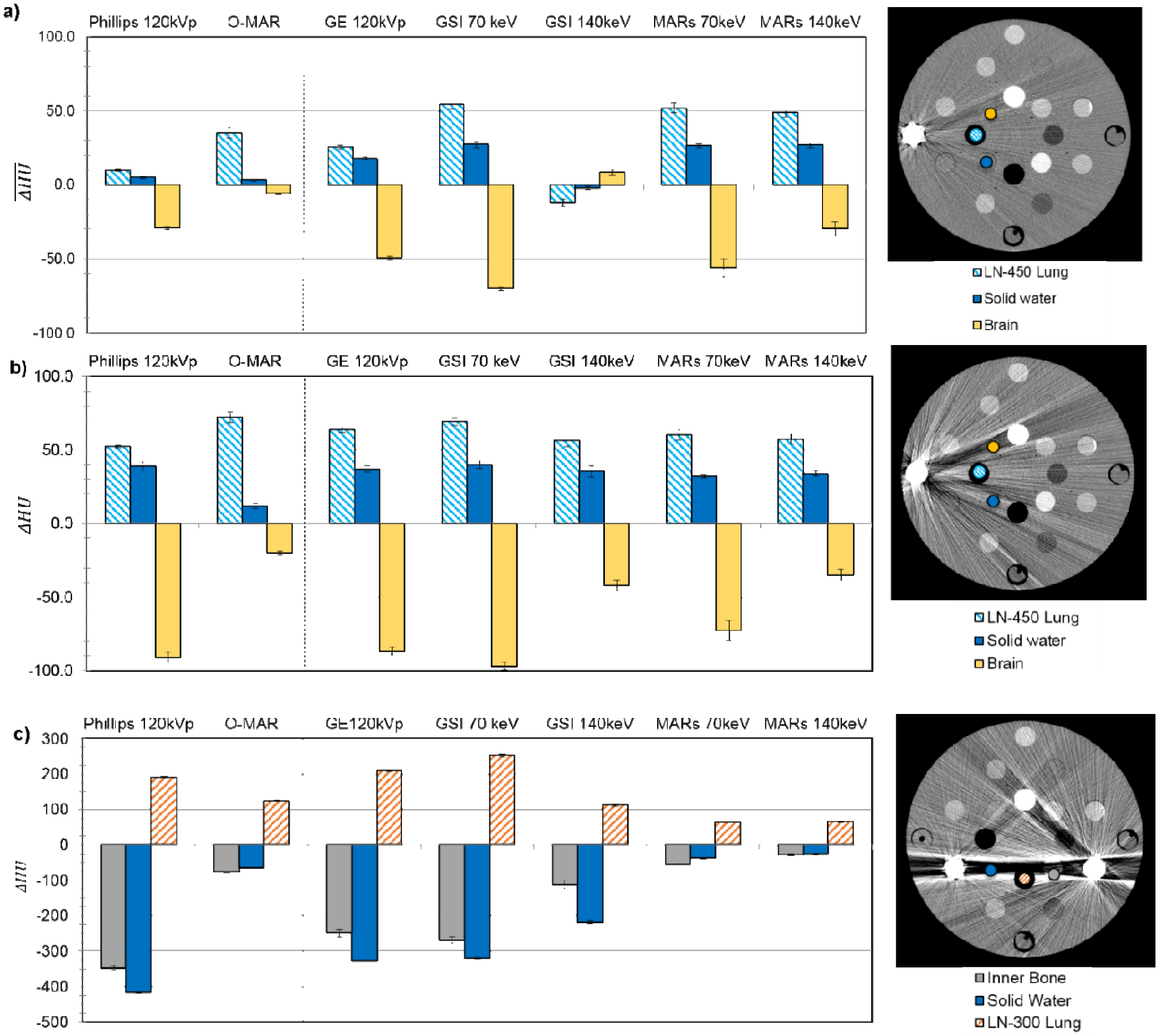


Figure 4.1: for select tissue substitue regions of interest in the RMI phantom scanned with (a) a unilateral titanium plug, (b) a unilateral stainless steel plug, and (c) bilateral stainless steel and titanium plugs. are grouped by imaging techniqe, including uncorrected imaging methods (120kVp) as well as the metal artifact reduction methods. For each plot, a CT image (Philips 120kVp protocol, WL=0, WW=500) on the right shows the location of the tissue substitute inserts for which is plotted and the position of metal inserts in the phantom. Error bars indicate the standard error of the mean for three repeated scans.

Table 4.2: Mean metal area, standard deviation of the metal area across five slices (σ), and calculated diameter for CT images of a stainless steel plug/rod scanned in the RMI phantom and titanium rods scanned in the anthropomorphic thoracic phantom. The diameter error is the difference between the calculated diameter and the physical diameter measured with electronic calipers, 28.6 mm for stainless steel and 9.5 mm for titanium.

	Stainless steel rod				Titanium rod			
Imaging technique	Metal area (mm^2)	σ (mm^2)	Calculated diameter (mm)	Diameter error (mm)	Metal area (mm^2)	σ (mm^2)	Calculated diameter (mm)	Diameter error (mm)
Philips 120kVp	704.0	2.1	29.9	1.4	169.4	0.5	10.4	0.9
O-MAR	693.2	1.6	29.7	1.1	166.7	0.4	10.3	0.8
GE 120kVp	695.0	1.6	29.7	1.2	168.2	0.5	10.3	0.9
GSI 70keV	701.2	0.9	29.9	1.3	175.5	0.7	10.6	1.1
GSI 140keV	699.5	1.2	29.8	1.3	148.2	0.9	9.7	0.2
MARs 70keV	582.1	2.3	27.2	-1.3	76.5	1.3	7.0	-2.5
MARs 140keV	578.2	0.7	27.1	-1.4	74.0	0.9	6.9	-2.6

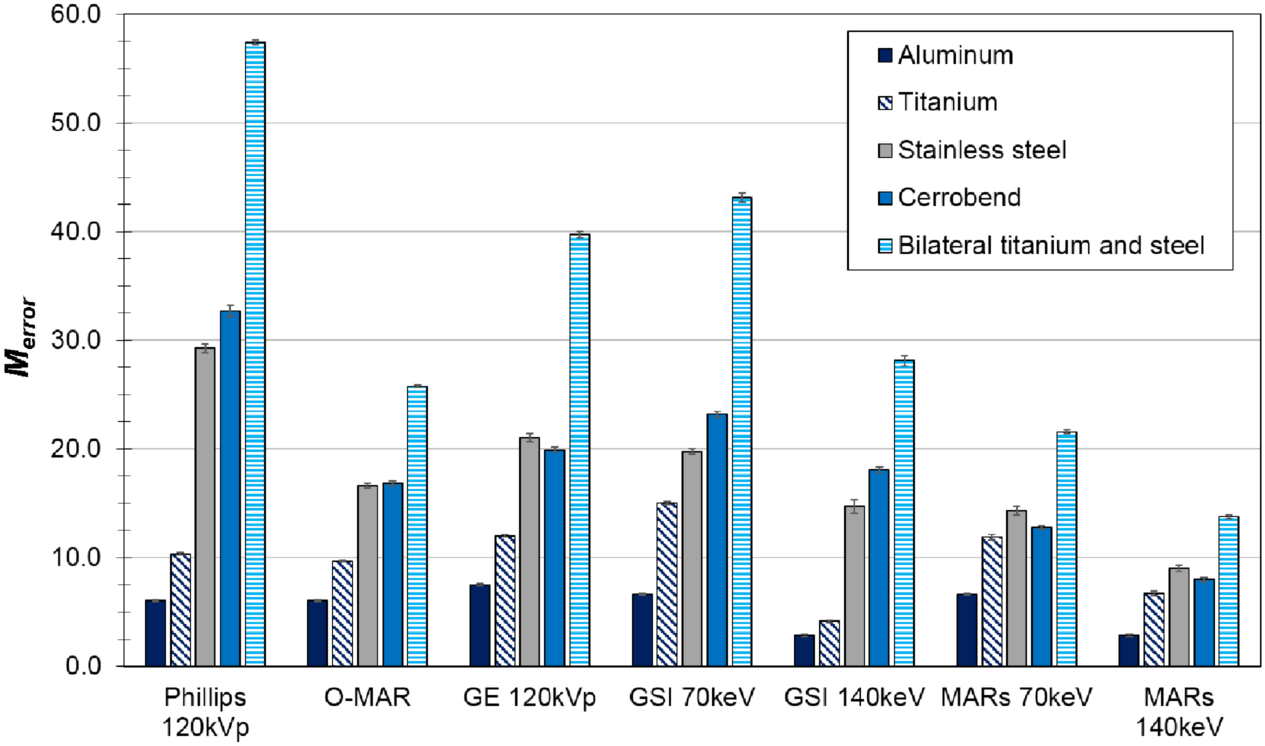


Figure 4.2: M_{error}, the fraction of bad pixels in the phantom image multiplied by the mean absolute CT number error of the bad pixels, for various imaging techniques and metal scan configurations of the RMI phantom.

Anthropomorphic phantoms

The anthropomorphic phantoms were evaluated qualitatively and quantitatively for their severity of streak artifacts. The quantitative results (% bad pixels and M_{error}) are summarized in Table 4.3.

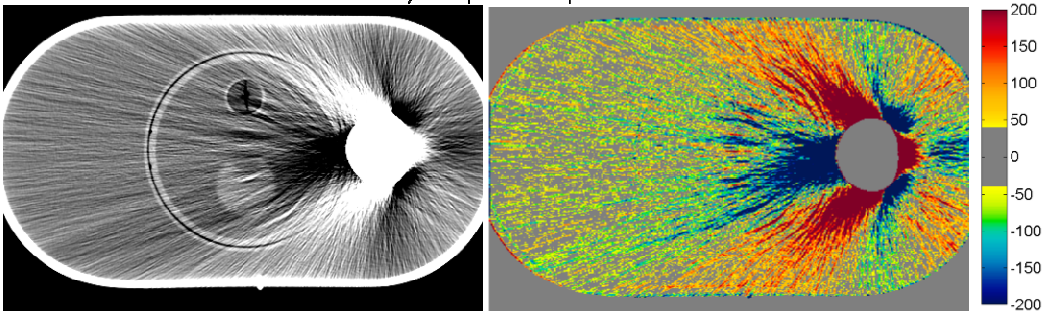
Table 4.3: Summary of quantitative results for three anthropomorphic phantoms with metal implants. Listed for each phantom are the number of images analyzed, the mean percentage of bad pixels for those images, and the error metric M_{error}, which takes into account both the number of bad pixels and the magnitude of CT number errors of the bad pixels.

Imaging technique	Hip prosthesis			Dental fillings			Spine stabilization rods		
	No. of images	% bad pixels	M_{error}	No. of images [a]	% bad pixels	M_{error}	No. of images	% bad pixels	M_{error}
Philips 120kVp	36	46.6	61.0	6	27.2	50.1	21	6.9	7.9
O-MAR	36	31.6	29.6	6	25.0	41.0	21	9.1	7.2
GE 120kVp	44	36.8	38.0	6	22.4	41.7	21	12.2	12.9
GSI 70keV	44	33.8	36.4	6	22.8	42.2	21	19.2	22.0
GSI 140keV	44	29.2	32.5	6	20.4	38.0	21	5.9	4.3
MARs 70keV	44	19.1	15.1	6	32.6	55.8	21	17.4	12.7
MARs 140keV	44	12.1	10.0	6	24.7	35.5	21	11.6	8.0

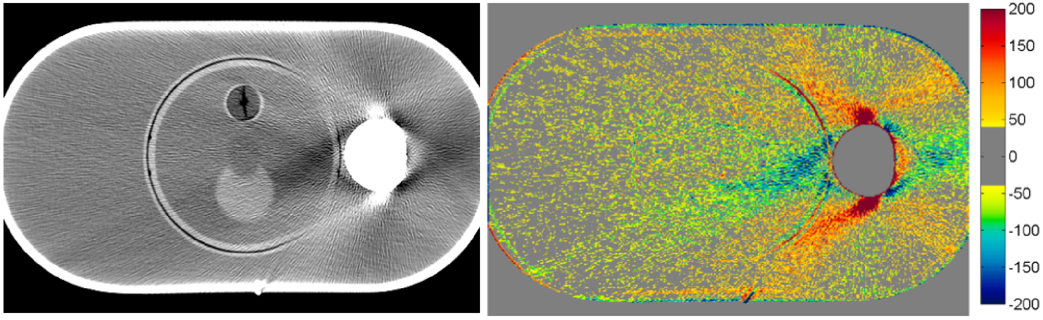
[a] Includes two images (one superior and one inferior to the images containing the metal fillings) that do not contain any metal

Pelvic phantom with hip prosthesis. The artifact mitigation methods were generally successful in reducing the severity of streak artifacts for the pelvic phantom with hip prosthesis. This can be seen qualitatively in Figure 4.3, which shows CT images and CT number difference maps between the metal-free baseline and metal (with Co-Cr hip prosthesis) images of this phantom. Both the number and severity of bad pixels were generally reduced by the artifact mitigation methods, as can be seen by a decrease in both the % of bad pixels and M_{error} compared with the uncorrected 120kVp images (Table 4.3). Philips O-MAR reduced the mean percentage of bad pixels from 46.6% to 31.6% and reduced M_{error} by a factor of 2. For the GE system, GSI 140keV imaging resulted in a slight improvement in the severity of streak artifacts, reducing the mean percentage of bad pixels from 36.8% to 29.2%, compared with conventional 120kVp imaging. MARs 140keV imaging gave the most substantial improvement, reducing the bad pixels further to 12.1% and reducing M_{error} by nearly a factor of 4 (Table 4.3). Based on these metrics, MARs 140keV was the most successful method for reducing the severity of streak artifacts for this phantom. However, it can be seen from Figures 4.3f and 4.3g that the MARs algorithm causes a decrease in the CT number in the center of the femoral head portion of the implant that was not observed for the other imaging methods.

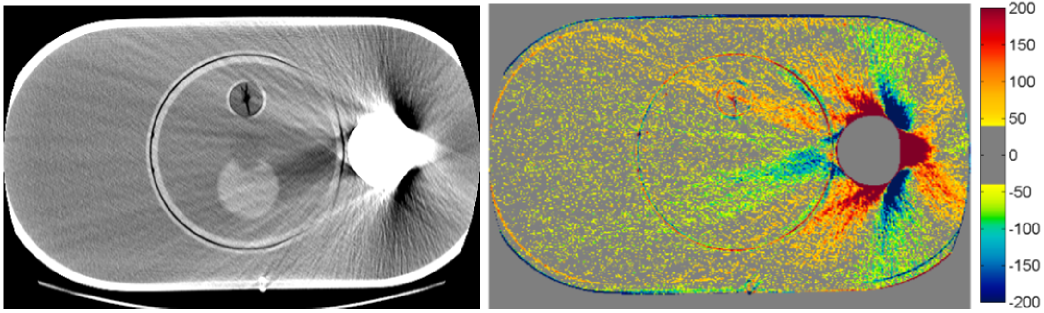
a) Philips 120kVp



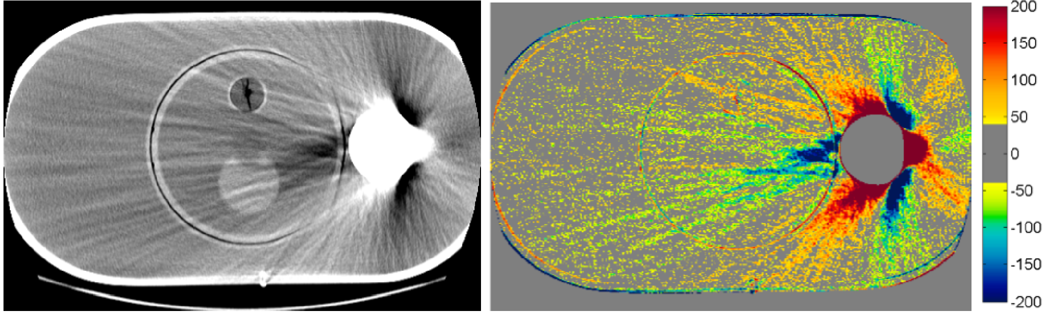
b) O-MAR



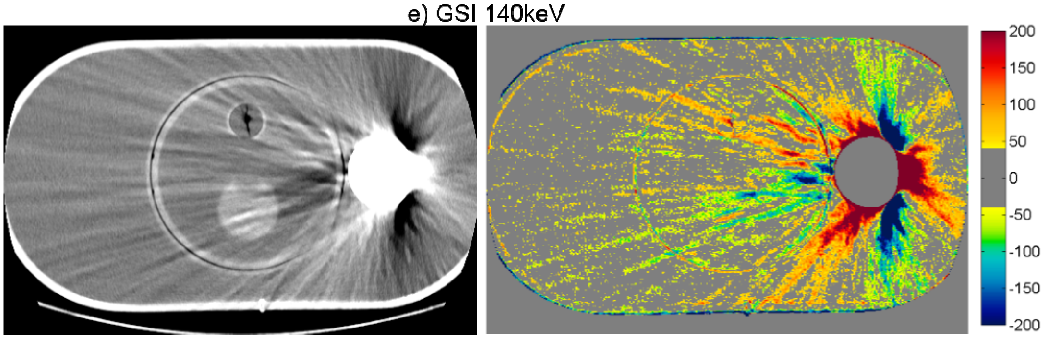
c) GE 120kVp



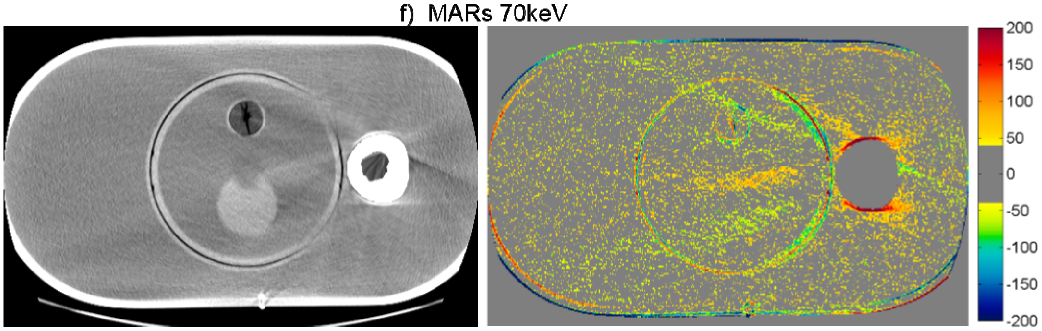
d) GSI 70keV



e) GSI 140keV



f) MARs 70keV



g) MARs 140keV

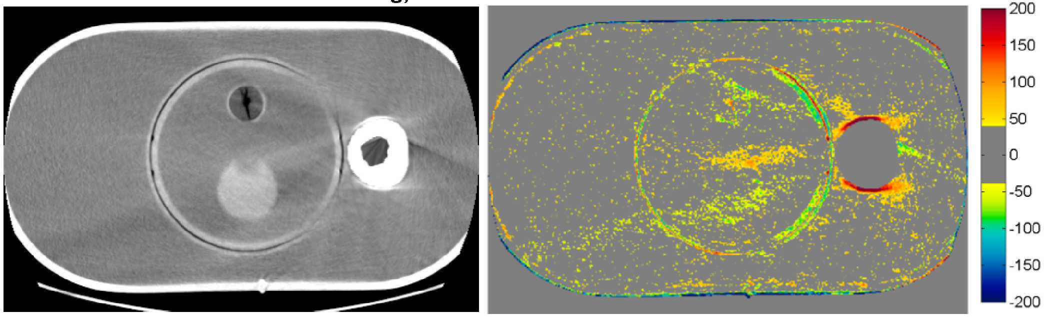
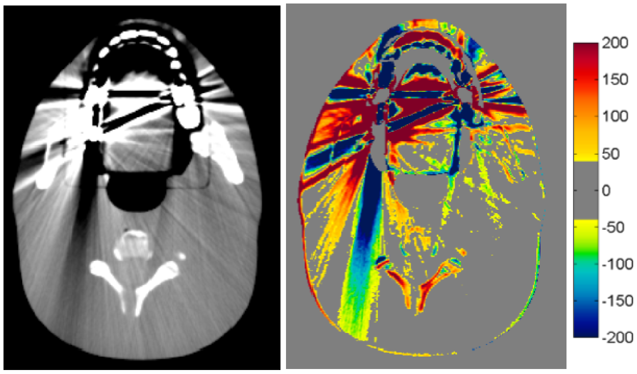


Figure 4.3. CT images of the pelvic phantom with hip prosthesis (WL=0, WW=500), side by side with the corresponding CT number difference maps between the baseline and the metal scans of the phantom for uncorrected imaging methods (“Philips 120kVp” and “GE 120kVp”) and artifact mitigation methods (“O-MAR”, “GSI”, and “MARs”).This image intersects the femoral head portion of the prosthesis.

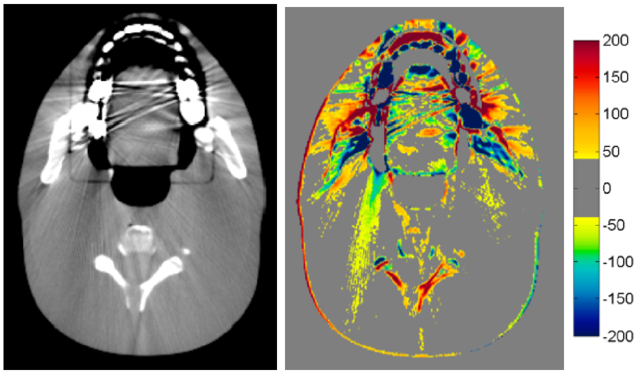
Head phantom with dental fillings. For the head phantom with dental fillings, none of the artifact mitigation methods were particularly successful, as is evident by the HU difference maps of Figure 4.4 and the quantitative results of Table 4.3. While Philips O-MAR and GSI 140keV imaging resulted in small reductions of the percentage of bad pixels (with a change of 2.2% and 2.0%, respectively), the MARs algorithm actually increased the percentage of bad pixels for both 70keV and 140keV imaging, meaning that the artifacts were worsened by the application of the MARs algorithm for this phantom. The image in Figure 4 shows some reduction in streak artifacts due to the artifact mitigation methods, specifically for O-MAR and

MARs 140keV. However, in other image planes, there was less success, particularly toward the edges of the fillings. In fact, the MARs algorithm introduced additional artifacts on images that contained no metal themselves but were adjacent to image locations containing the metal fillings ("out-of-plane" artifacts). Figure 4.5 illustrates these artifacts introduced by the MARs algorithm. Because this particular image contains no portion of the metal fillings, the GSI 140keV image contains very few HU errors (Figure 4.5a). However, MARs introduced artifacts in this image and caused a large increase in the number of bad pixels (HU error > 40) (Figure 4.5b). To take into account these "out-of-plane" artifacts in our artifact severity metrics, two additional images, one superior and one inferior to the metal fillings, were included in our data analysis (Table 4.3).

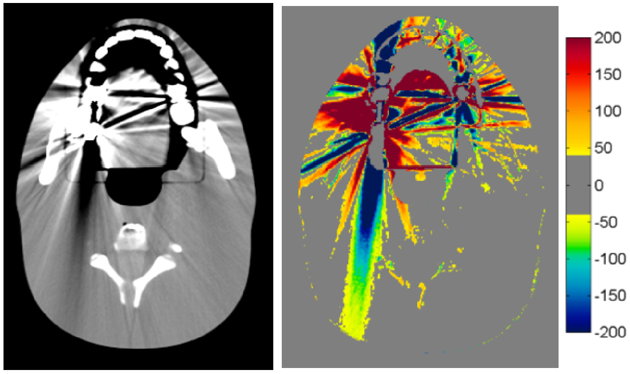
a) Philips 120kVp



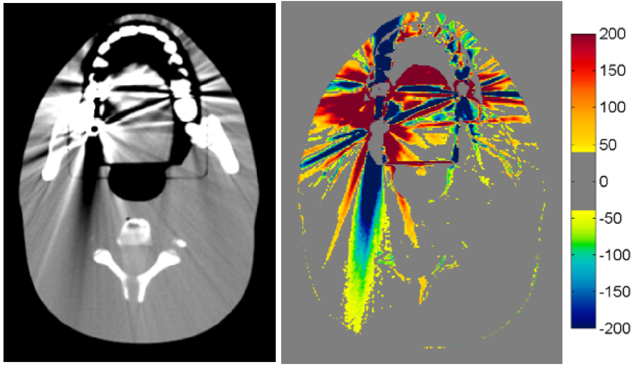
b) O-MAR



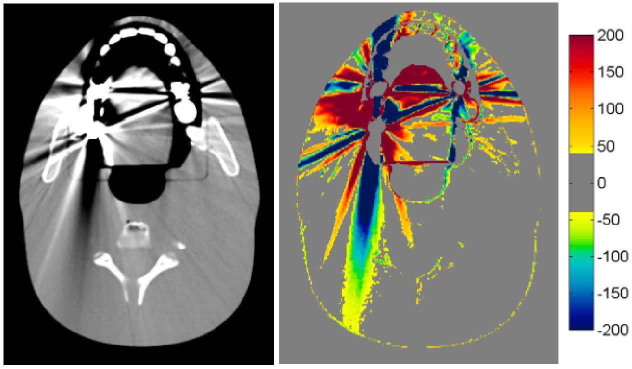
c) GE 120kVp



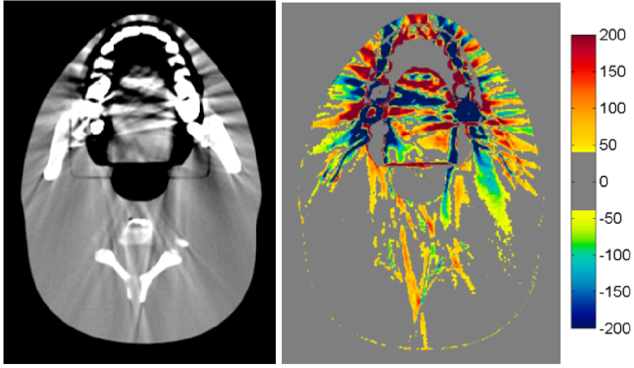
d) GSI 70keV



e) GSI 140keV



f) MARs 70keV



g) MARs 140keV

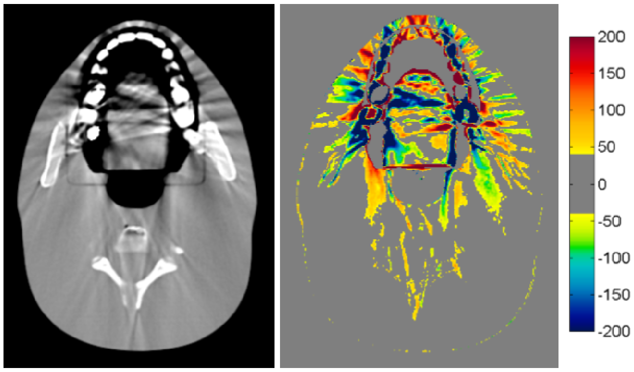
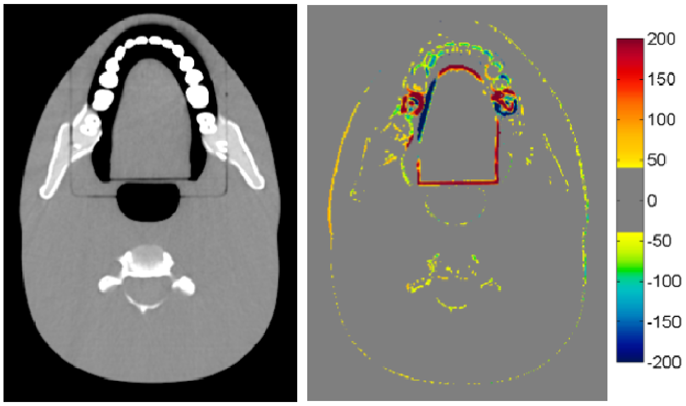


Figure 4.4: CT images of the head phantom with dental fillings (WL=0, WW=500), side by side with the corresponding CT number difference maps between the baseline and the metal scans of the phantom for uncorrected imaging methods (“Philips 120kVp” and “GE 120kVp”) and artifact mitigation methods (“O-MAR”, “GSI”, and “MARs”).

a) GSI 140keV



b) MARs 140keV

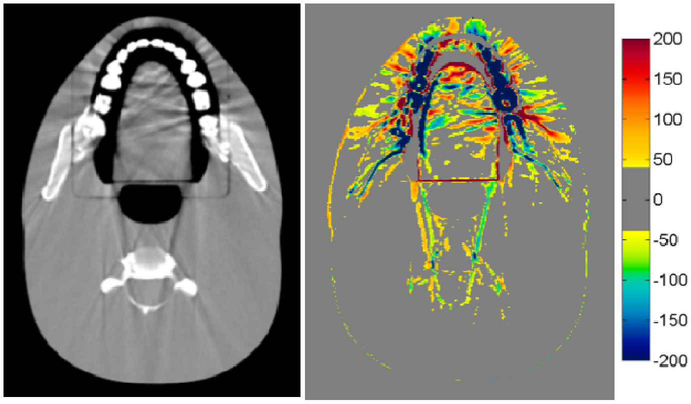
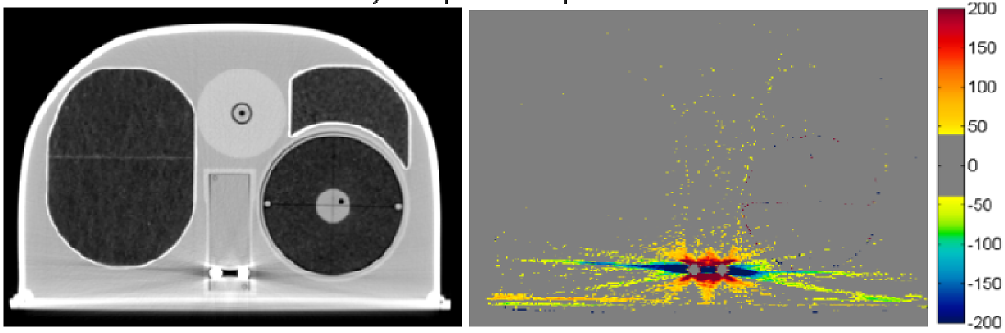


Figure 4.5: Grayscale CT images of the head phantom with dental fillings (WL=0, WW=500), side by side with the corresponding CT number difference maps between the baseline and the metal scans of the phantom for a) GSI 140keV and b) MARs 140keV imaging. Shown is an image of the head phantom that does not contain any portion of the metal fillings, illustrating out-of-plane artifacts introduced by the MARs algorithm.

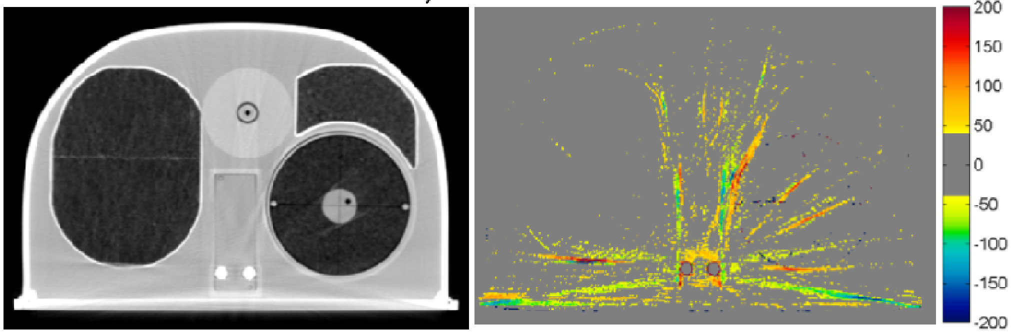
Thoracic phantom with spine stabilization rods. The qualitative ability of the methods investigated to mitigate CT artifacts for the thoracic phantom with spinal rods is shown in

Figure 4.6. O-MAR actually resulted in an increase in the percentage of bad pixels, introducing new artifacts farther away from the metal (Figure 4.6a versus 4.6b). However, this was offset by a decrease in the magnitude of HU errors of these bad pixels, particularly for pixels in close proximity to the metal rods, resulting in comparable values for the error metric M_{error} for imaging with and without O-MAR (Table 4.3). GSI and GSI with MARs also showed mixed results. At 70 keV, both GSI and GSI with MARs showed an increase in the percentage of bad pixels, although the overall severity (M_{error}) was reduced slightly for the MARs case. However, at 140 keV, both the percentage of bad pixels and the severity of the artifacts were improved with GSI and GSI with MARs. Of note, similar to the Philips O-MAR algorithm, the MARs algorithm introduced artifacts far away from the rods for both energies investigated (Figure 4.6f and 4.6g). GSI 140keV imaging (without MARs) was the most successful in reducing the severity of streak artifacts for this phantom and exhibited no added artifacts.

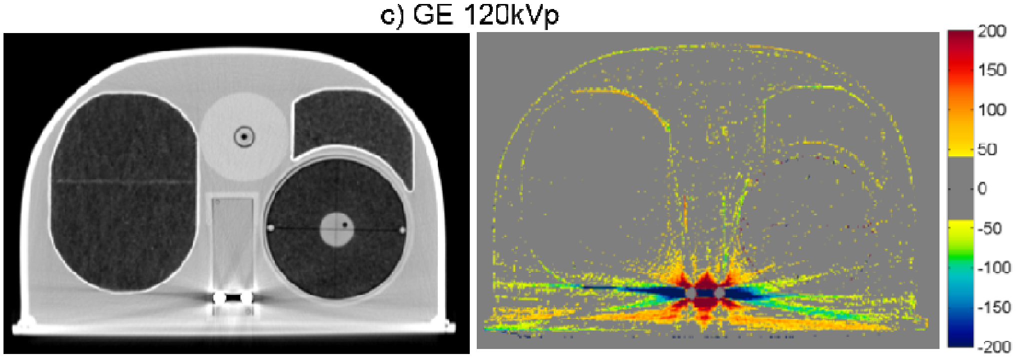
a) Philips 120kVp



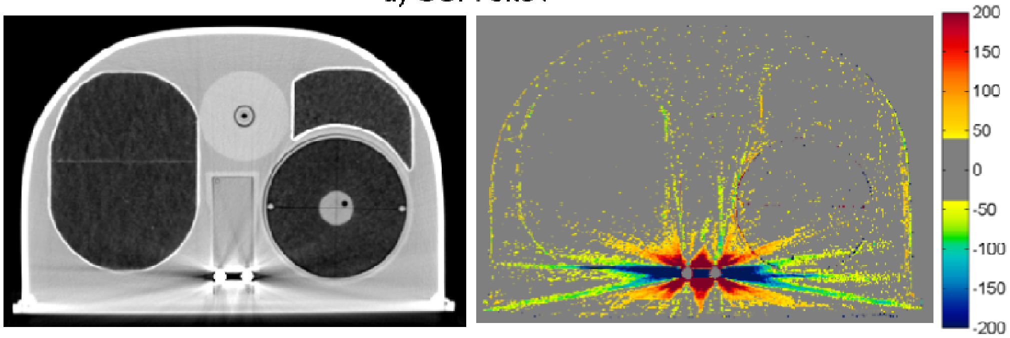
b) O-MAR



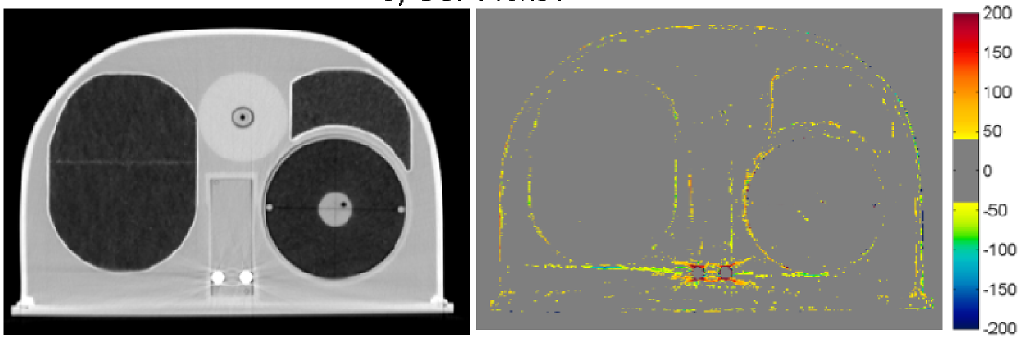
c) GE 120kVp

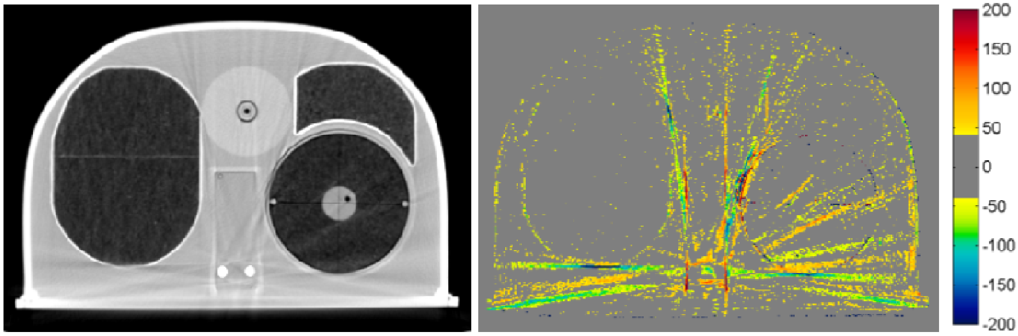


d) GSI 70keV



e) GSI 140keV




f) MARs 70keV

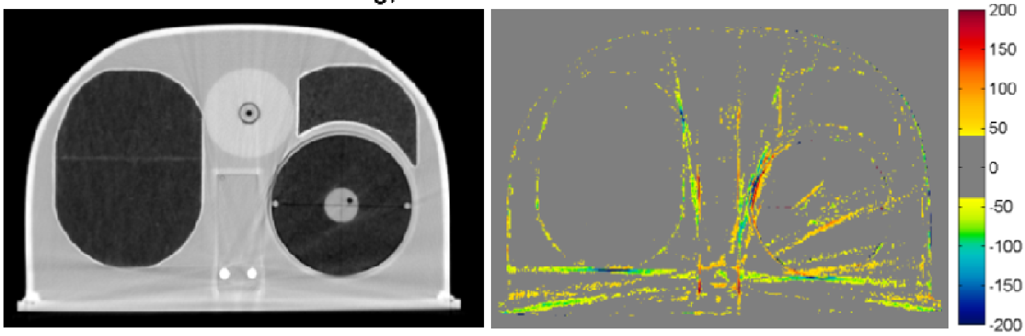

g) MARs 140keV

Figure 4.6: Grayscale CT images of the anthropomorphic thoracic phantom with titanium spinal rods (WL = -250, WW = 1250), side by side with the corresponding CT number difference maps between the baseline and the metal scans of the phantom for uncorrected imaging methods (“Philips 120kVp” and “GE 120kVp”) and artifact mitigation methods (“O-MAR”, “GSI”, and “MARs”).

The accuracy of the size of the titanium rods was also investigated by using the CT images of this phantom. Both O-MAR and GSI imaging were able to accurately represent the diameter to within 1.1 mm (approximately 1 pixel), as shown in Table 4.2. However, MARs imaging at 70kev and 140keV underestimated the diameter of the titanium rod by about 2.6 mm, a 26% underestimation. To further investigate this underestimation of the titanium rod diameter by the MARs algorithm, profiles were taken through the titanium rod for all imaging methods (Figure 4.7). From Figure 4.7, it can be seen that while GSI imaging and application of O-MAR do not affect the shape or the FWHM of the profiles, application of the MARs algorithm not only decreases the FWHM but also results in a less steep fall off at the edges of the titanium rods.

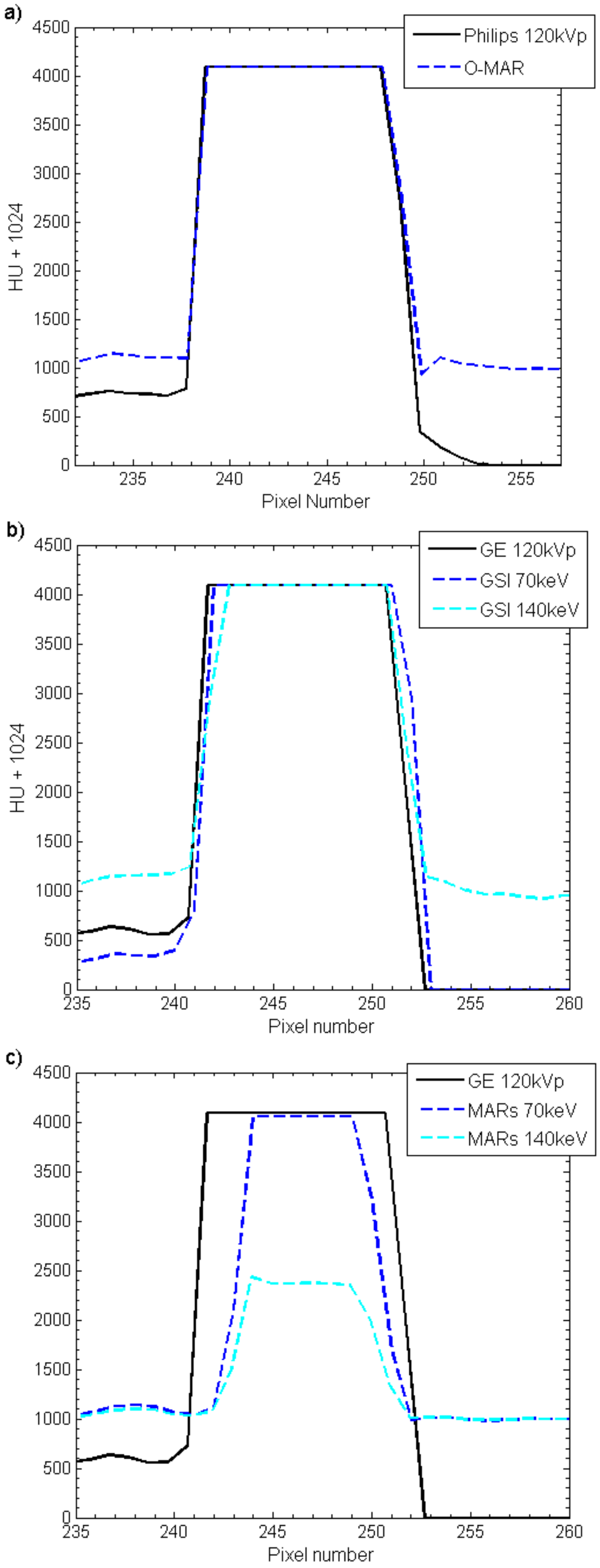


Figure 4.7: Horizonal pixel intensity profiles taken across one of the titanium rods scanned with the thoracic phantom for a) O-MAR, b) GSI imaging ("GSI 70keV" and "GSI 140keV"), and c) GSI imaging with MARs applied ("MARs 70keV" and "MARs 140keV"). The corresponding uncorrected imaging methods ("Philips 120kVp" and "GE 120kVp") are also shown for comparison.

4.4 Discussion

In this study, three commercial metal artifact reduction methods were evaluated using metrics that evaluate image quality in the context of radiation therapy. Three anthropomorphic phantoms were used to evaluate how successful each method was at reducing artifacts for three common types of metal implants. Although each method exhibited some success in improving CT images, none of the methods were globally effective for all of the sites investigated, and some exhibited some limitations that users should be aware of. Strengths and weaknesses of each method are summarized in Table 4.4.

Table 4.4: Summary of our results for the various anthropomorphic phantoms and general impressions of the metal artifact reduction methods. ✓ indicates that the method resulted in a small reduction in streak artifacts, ✓✓ indicates a more substantial reduction in streak artifacts, and ★ indicates that the method was the most successful method of the three investigated and was highly successful for the given site. X indicates that the method is not recommended for use at a particular site.

	Pelvic	Head	Thoracic	Weaknesses/Drawbacks
O-MAR	✓✓	✓	✓	• Induced artifacts for thoracic phantom
GSI 140keV monochromatic imaging	✓	✓	★	• No major drawbacks identified
MARs 140keV monochromatic imaging	★	X	✓	• Underestimation of metal size and possible distortion of metal shape • Induced “out-of-plane” artifacts for dental fillings • Induced artifacts for thoracic phantom

For the case of a unilateral hip prosthesis, both O-MAR and MARs were fairly successful in reducing the severe artifacts caused by the implant. The geometry of the hip prosthesis case represents perhaps the most ideal geometry for metal artifact reduction methods in that a large metal implant is located in a fairly homogeneous environment. Thus, it is not surprising that the metal artifact reduction methods were successful for this particular site. In fact, O-MAR is designed primarily for orthopedic implants, and Li *et al.* (2012) previously found that O-MAR was successful for patients with hip prostheses, particularly for patients with large bilateral hip prostheses. The MARs algorithm is similar to O-MAR in that it is a projection modification approach to metal artifact reduction, whereby the projections affected by the metal object are identified and corrected for based on unaffected data. However, MARs has the advantage that it is applied onto virtual monochromatic images that exhibit reduced beam hardening artifacts in comparison to polyenergetic imaging. Thus, the MARs 140keV imaging was the most successful method for the case of the hip prosthesis.

While the hip prosthesis case was an ideal case for metal artifact reduction, the case of dental fillings is perhaps the most challenging geometry. Dental fillings are small metal implants surrounded by a highly heterogeneous local environment including the teeth as well as air cavities. The methods investigated in this study were generally not successful in reducing the artifacts caused by dental fillings. Despite being designed for orthopedic implants, O-MAR did offer a slight benefit for the dental fillings, as did GSI virtual monochromatic imaging at 140keV, although gains were fairly modest. Notably, application of the MARs algorithm to the monochromatic images resulted in an increase in the overall severity of streak artifacts. Furthermore, MARs caused "out-of-plane" artifacts in adjacent image locations that did not contain any metal (Figure 4.5). Although the exact cause of these "out-of-plane" artifacts is not known to the user, it is postulated that the MARs algorithm is sensitive the heterogeneities in close proximity to the metal implant, e.g., teeth and air gaps. The fact that induced artifacts are observed in adjacent image plans suggests that MARs performs some level of smoothing on the sinogram data, perhaps with a 3D convolution kernel that may not be appropriate for small metal implants such as dental fillings.

The case of the anthropomorphic thoracic phantom with two titanium spinal rods represents an intermediate level of difficulty for the metal artifact reduction methods in that both the size of the metal implant and the level of heterogeneity in the environment surrounding the metal are intermediate between the case of the hip prosthesis and that of the dental fillings. For this phantom, we found that both O-MAR and MARs introduced similar artifacts; streaks were introduced between the titanium rods and the edges of the heart structure and between the rods and the edges of the target in the left lung. Interestingly, the artifacts were not introduced in a symmetric manner since no additional streaks were observed in the right lung, suggesting that these projection modification algorithms struggle with heterogeneities and material interfaces. Brook *et al.* (2012) also observed a similar introduction of additional artifacts when MARs was used for patients with gold fiducial markers, and other studies evaluating metal artifact reduction methods that perform linear interpolation of projection data also observed secondary artifacts introduced between metals and heterogeneities such as bone and contrast material.[19, 64] It should be noted that Philips states that O-MAR is contraindicated for cases in which a metal implant is located near low-density tissue, such as lung, although the specifics of this recommendation are vague in terms of proximity.[36] For this phantom, GSI monochromatic imaging was the most successful method for reducing artifacts caused by the titanium rods with 140keV virtual monochromatic images showing nearly complete artifact reduction (Figure 4.6e).

GSI dual-energy CT data can be reconstructed at any energy from 40keV to 140keV to generate virtual monochromatic images. In this study, we evaluated 70keV and 140keV only. One limitation of this study is that only two energies were investigated. Lee *et al*. (2012) investigated MARs for metal artifact reduction of titanium and steel implants and found that 80keV and 110keV were the optimal energies for titanium and stainless steel respectively, while Wang *et al*. (2013) found that the optimal monochromatic energy level for pedicle screws was 110-140keV. Thus, the two energies we chose to investigate spanned the energies found to be successful in the literature. In our phantom studies, we found that GSI monochromatic 70keV images gave similar results to 120kVp images, while monochromatic 140keV images showed better artifact reduction than 70keV. The reduced artifacts in the 140keV images in comparison to the 70keV images can be explained by the fact that this high energy reconstruction has a higher proportion of information from the high energy projection data vs. the low energy projection data (140kVp vs. 80kVp). Since the high energy projections contain reduced beam hardening in comparison to the low energy projections, this reduced beam hardening propagates to the 140keV virtual monochromatic images.

We also investigated how accurately the various metal artifact reduction methods were able to represent the size of metal objects. We found that MARs underestimated the diameter of both stainless steel and titanium rods. For 29 mm diameter stainless steel rod, MARs images were able to preserve the diameter to within 1.4 mm (5% underestimation). However, for 9.5 mm diameter titanium rods, application of MARs resulted in a 2.6 mm (26%) underestimation of the diameter (Table 4.2). In agreement with our results, Lee *et al.* (2012) found that while 140keV MARs images gave accurate dimensions for a stainless steel prosthesis (within 1 mm), the thickness of titanium was underestimated by approximately 3 mm. Wang *et al.* (2013) investigated MARs for patients with pedicle screws and found that the MARs algorithm resulted in unacceptable distortion in the shape and size screws. Profiles through the titanium rod revealed that not only did MARs decrease the width of the metal profile but MARs also affected the shape of the profile (Figure 4.7c). In comparison to the profiles acquired from GSI images, the MARs profiles exhibit a larger penumbra region, suggesting that some sort of smoothing was applied with the MARs algorithm. It can also be seen from the MARs 140keV profile in Figure 4.7c that MARs affects the HU values of metal implants. We investigated how MARs affects the HU values of various metals in our study (titanium, stainless steel, and Co-Cr alloy) and found that MARs consistently maps metals to a pre-defined HU level, with these pre-defined HU values varying as a function of monochromatic reconstruction energy. This data suggests that metal pixels are identified prior to application of the MARs algorithm, and these pre-defined metal HU values are inserted back into the image after application of MARs. Interestingly, it was also observed that the MARs algorithm can decrease HU values in the center of large metal implants (Figures 4.3f and 4.3g).

In addition to metal objects appearing smaller than reality in the image plane, we also observed a distortion in the size of metal objects along the scan direction. Specifically, for the hip prosthesis, no metal was visible on the MARs image at the image location containing the most superior portion of the femoral head component of the prosthesis (this metal was visible at the same image location with both GSI imaging without MARs and 120kVp polyenergetic imaging). The same effect was observed for the edge of the titanium rods in the spine phantom. Brook *et al.* (2012) found similar distortions, in that MARs caused some gold fiducial markers to be barely visible in patient scans. These distortions again suggest that MARs performs some form of smoothing on the image data.

Successful reduction of streak artifacts will allow more confidence in the contouring of the target and surrounding structures, allow more flexibility of beam arrangements, and improve dose calculation accuracy by providing more accurate CT numbers. Li *et al.* (2012) found small dosimetric differences (generally <1% of prescription dose) between treatment plans calculated on the O-MAR vs. non-OMAR images for prostate cancer patients with unilateral hip prostheses. However, in the head and neck region, where targets and critical organs can be located very close to dental restorations, the artifact mitigation methods may have a greater impact on dose calculation accuracy. Future studies are planned to investigate the dosimetric impact of these metal artifact reduction methods, in conjunction with a novel implementation of collapsed cone convolution/superposition dose calculation using metal kernels.[65]

4.5 Conclusions

Commercial metal artifact reduction methods were evaluated for their effectiveness in reducing metal artifacts in CT images. Although the metal artifact reduction methods were evaluated based on metrics that are most relevant to treatment planning dose calculation accuracy, our results nonetheless provide useful information about CT imaging in general. Our data suggest that all three of the artifact mitigation methods can be used effectively for large orthopedic implants in fairly homogenous environments, such as for hip prostheses, while more varied results were observed for artifacts caused by small metal implants in heterogeneous environments, such as dental fillings or spinal fixation rods. Both the O-MAR and MARs algorithms introduced secondary artifacts when applied in the heterogeneous environment of the thorax. The MARs algorithm should be used with caution in certain scenarios, as it was found to underestimate the size of metal implants and introduced new artifacts into imaging planes beyond the metal when applied to dental artifacts. Although GSI virtual monochromatic imaging was not observed to cause any additional artifacts in our phantom studies, it did not offer as much of a benefit as the other two methods for large orthopedic implants.

Chapter 5: Dosimetric impact of implementing metal artifact reduction methods and metal energy deposition kernels for photon dose calculations

5.1 Introduction

Many patients receiving external beam radiation therapy have metal implants that can complicate the treatment process. In a survey conducted by the American Association of Physicists in Medicine's Task Group 63, it was found that 1-4% of all radiation therapy patients have a prosthetic device that could affect their treatment.[1] This percentage does not include the large number of patients with dental fillings that affect the treatment of head and neck cancers. The presence of metal implants poses certain challenges in the treatment planning process and negatively impacts the dose calculation accuracy of the resulting treatment plan. Dose calculation errors associated with metal implants stem from computed tomography (CT) imaging artifacts, as well as limitations in modern dose calculation algorithms.

Metal implants cause well-known imaging artifacts in the CT images that are used for treatment planning.[14] These metal streak artifacts make it difficult to confidently delineate the tumor and surrounding organs and can also negatively impact dose calculation accuracy. Imaging artifacts result in errors in CT numbers, which propagate to density assignment errors and subsequently to dose calculation errors.[15, 16] One strategy to mitigate dose calculation errors associated with CT imaging artifacts is to use a metal artifact reduction method. Recently, commercial metal artifact reduction methods have become available. One of these commercial solutions, the algorithm for orthopedic implants (O-MAR) developed by Philips Healthcare (Cleveland, OH) is becoming increasingly popular for CT simulation imaging in radiation oncology clinics. The O-MAR algorithm is an iterative projection modification solution that identifies projection data corrupted by the presence of the metal implant and corrects it based on nearby uncorrupted data.[36] Another approach to metal artifact reduction is the use of dual-energy CT. Dual energy CT has many applications for diagnostic imaging, [38-40] but few studies have looked at its use for treatment planning.[61, 66] One such dual-energy CT system, GE Healthcare's Discovery CT750 HD (Milwaukee, WI), acquires dual-energy projection data via fast kilovoltage switching with a single X-ray source. This dual-energy projection data can then be reconstructed to generate virtual monochromatic images at various energy levels (from 40keV to 140keV), called Gemstone Spectral Imaging (GSI). These monochromatic images depict how an object would look if it were imaged using a monoenergetic X-ray source[42] and have reduced beam hardening artifacts in comparison to conventional polyenergetic images.[38, 40, 60] To further reduce artifacts, GE has developed metal artifact reduction software (MARs) specifically for use with GSI monochromatic imaging that reduces artifacts caused by photon starvation.[43]

In addition to errors associated with imaging artifacts, errors also result from the limited ability of modern dose calculation algorithms to accurately model radiation transport in and near metal implants. The convolution/superposition (C/S) method is the current standard of care in commercial treatment planning systems (TPS) for photon dose calculations. In the C/S method, the energy released by photon interactions in the patient is described by the TERMA (total energy released per unit mass), while the process of energy deposition via secondary electrons and scattered photons is described by the energy deposition kernel. The TERMA and energy deposition kernel are convolved to calculate the absorbed dose. The kernels used in commercial TPSs are calculated using Monte Carlo simulations based on photon interactions and scatter in water.[29] Although C/S is generally able to calculate the attenuation caused by metal implants in ideal cases (in which the density and physical dimensions of metal implants are known), C/S inaccurately models tissue/metal interface effects. Studies investigating the accuracy of C/S for common implant materials have found that C/S underestimates the backscatter dose enhancement upstream of the metal and overestimates the dose downstream of the metal.[3, 10, 11]

One source of this dose calculation error is that for heterogeneous dose calculations, the density scaling approximation is used, in which the water-based kernels are simply scaled in dimension based on the local density.[33, 34] This density scaling of water kernels is only a good approximation for materials with the same atomic composition as water and thus may be inadequate to describe the physical interactions and scatter occurring in metal implants. Recently, Huang et al.[65] generated kernels for several metals and found that the shape of the dose deposited for these metal kernels differed from that of water kernels, suggesting that density scaling is not a good approximation for metals. Furthermore, they found that implementing metal kernels in a commercial C/S algorithm resulted in increased backscatter dose and decreased dose downstream of metal implants, which suggests that implementation of metal kernels may correct the interface errors associated with conventional C/S algorithms that perform density scaling of water kernels.

Although some treatment planning studies have been performed to evaluate the O-MAR algorithm,[37, 58, 67] very little has been published on the use of GSI dual-energy CT imaging for treatment planning,[61, 68] and no studies have investigated the impact of GSI dual-energy CT imaging or the GSI-MARs algorithm on the accuracy of dose calculations. A thorough study investigating the dosimetric impact of these metal artifact reduction methods would allow comparison between the different methods and give information relevant to the care of patients with metal implants. Furthermore, although Huang et al.[65] investigated how implementing metal kernels affects the dose in comparison to density scaling of water kernels, no reported work has validated that implementing metal kernels improves calculation accuracy near metal/tissue interfaces, or quantified the magnitude of any such improvement. Thus, the purpose of the study is to evaluate the dosimetric impact of the use of commercial metal artifact reduction methods for CT imaging and the use of metal kernels in the C/S dose calculation method. Both error-reduction strategies were evaluated with a geometric slab phantom with two different metals to mimic both dental and nondental implants.

5.2 Methods and Materials

Phantom geometry

Both the metal artifact reduction methods and the metal kernel implementation were evaluated using a simple slab phantom. This geometric phantom was composed of 30 cm x 30 cm high-impact polyethylene slabs with a rectangular cavity that could accommodate a metal insert. The phantom also had space for film to measure the dose upstream and downstream of the metal (Figure 5.1). Two different metals, titanium (4.51 g/cm^3) and Cerrobend (9.4 g/cm^3), were used with this phantom. Titanium was chosen because it is a common implant material for both hip prostheses[1] and spinal fixation devices,[3, 4, 69, 70] and Cerrobend was chosen because it has a relative electron density similar to that of dental amalgam.[10] The lateral dimensions of both metal inserts were 7.5 cm x 7.5 cm, while the thickness of the titanium and Cerrobend inserts were 2 cm and 4 mm, respectively. To measure the dose as close as possible to the metal/tissue interfaces and to ensure high spatial resolution, we placed stacks of 3 cm x 3 cm EBT2 radiochromic films (Gafchromic, Ashland, Wayne, NJ), totaling 5 mm in thickness and oriented perpendicular to the central axis of the radiation beam, at both the proximal and distal interfaces. To measure the dose further away from metal/tissue interfaces, we placed films upstream and downstream of the stacked films on the central axis, oriented parallel to the beam direction (Figure 5.1).

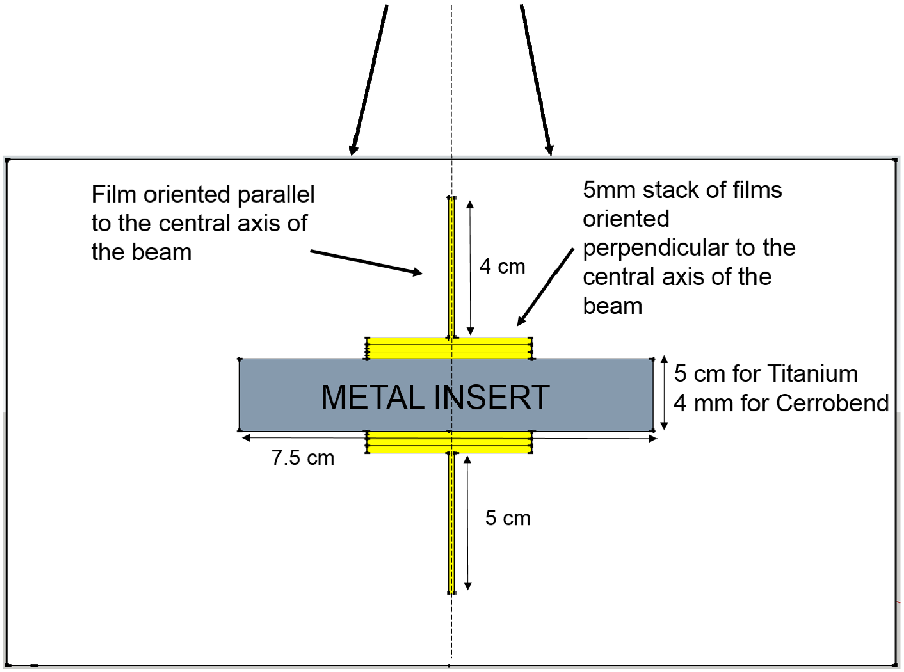


Figure 5.1: Schematic of the measurement phantom used for this study, showing the location of the metal insert along the central axis and the locations of films used to measure the dose upstream and downstream of the metal.

Phantom irradiation and film analysis

The phantom was irradiated using a 6-MV photon beam from a Varian Clinac 2100 linear accelerator with a 5cm x 5cm field size and a 100-cm source to surface distance. For each irradiation, all of the films were placed in the phantom so that a depth dose curve could be generated. For each of the two metal implants, measurements were repeated three times to create an average percent depth dose curve.

The optical density (OD) of the irradiated films was obtained using a CCD Microdensitometer for Radiochromic Film model CCD100 (Photoelectron Corp., Lexington, MA) approximately 4 days after irradiation. For the interface films oriented perpendicular to the beam, ImageJ (National Institute of Health, Bethesda, MD) was used to obtain the mean OD value for a 30 pixel x 30 pixel region of interest (0.76 cm^2) centered on the central axis. For the films oriented parallel to the beam, ImageJ was used to obtain average OD values (for a 30-pixel-wide region of interest centered on the central axis) as a function of depth. A background

OD value, measured from an unirradiated film of the same batch, was subtracted from all irradiated films. In order to convert the net OD to absorbed dose, we generated a calibration curve (third-degree polynomial fit) by irradiating films to absorbed doses ranging from 0.5 Gy to 15 Gy. The measured dose from the parallel and perpendicular films were then combined to create a depth dose curve. Because the dose response for the films differed according to their orientation relative to the beam direction,[71] different normalization factors were used for the parallel and perpendicular films. The parallel films were normalized using the maximum measured dose at d_{max}. The perpendicular interface films were normalized such that a continuous depth dose curve was obtained, i.e., such that there were no discontinuities where the perpendicular films met the parallel films.

M3D dose calculations

To quantify dose calculation accuracy, we compared the measured dose with dose calculations performed using Mobius3D v1.3.1 (Mobius Medical Systems, Houston, TX), a commercial TPS verification tool that calculates dose using a collapsed cone C/S algorithm similar to Pinnacle's collapsed cone algorithm (CCC).[25] Mobius3D (M3D) utilizes reference or stock beam models for existing models of linear accelerators. To better match the specific Varian Clinac 2100 accelerator used for the film measurements, we customized the stock beam model by inputting several dosimetry parameters (percent depth dose values for different depths and field sizes, off-axis factors, and output factors) into M3D's automodeling feature. This customized beam model was then validated with various verification treatment plans, ranging in complexity from a simple water phantom with a single treatment beam to intensity-modulated radiotherapy (IMRT) clinical plans. Dose calculations from the final M3D beam model agreed very well with those from our institution's clinical Pinnacle treatment planning system (Philips

Healthcare, Andover, MA), with all of our verification plans resulting in > 96% of pixels passing for 3D gamma analysis at ±3%/3mm (Appendix C).

For heterogeneity corrections, CT number-to-density conversion curves were created for the M3D system using the RMI 467 tissue characterization phantom (Gammex, Middleton, WI), which contained several tissue substitute materials of known density, including adipose, cortical bone, and lung. Two different curves, shown in Figure 5.2, were created, one for conventional polyenergetic CT imaging ("120kVp" curve) and one for GSI monochromatic imaging at 140keV with and without MARs ("140keV" curve). In agreement with Yagi et al.,[61] the 120kVp curve is bilinear, while the 140keV curve is fairly linear and has a steeper slope. Because this study sought to assess the improvement in dose calculation accuracy due to the application of different metal artifact reduction methods, we wanted the density mapping of the metal implants to be consistent across the different imaging techniques regardless of the different slopes in the two conversion curves. Therefore, for all dose calculations, the saturated Hounsfield unit (HU) values in the images were mapped to the known density of the metals, essentially performing a density override for the metal region in the CT images. For our 12-bit CT images, the maximum HU value was 3071. However, the MARs algorithm consistently maps metals to a pre-defined HU level that is a function of the monochromatic reconstruction energy.[68] For 140keV GSI imaging with MARs, this value was approximately 1340 HU, so HU values greater than or equal to this threshold value in the MARs images were assigned to the true metal density.

To assess how metal kernels might impact dose calculation accuracy, we implemented metal kernels into M3D's collapsed cone C/S algorithm. For dose calculations with the titanium implant, titanium kernels were used, and for dose calculations with the Cerrobend implant, silver kernels were used. These metal energy deposition kernels were simulated using the EGSsnrc Monte Carlo system, characterized in previous work,[65] and implemented in the M3D algorithm using a density threshold. For voxels with an assigned density greater than the threshold value, metal kernels were used to describe the energy deposition for energy released from those voxels. For voxels with assigned density less than the metal threshold, water kernels were used. The density threshold value was chosen such that voxels assigned the true metal density would use the metal kernels. Kernel density scaling was still employed for voxels with a density that did not exactly match one of the two kernel materials (water or metal).

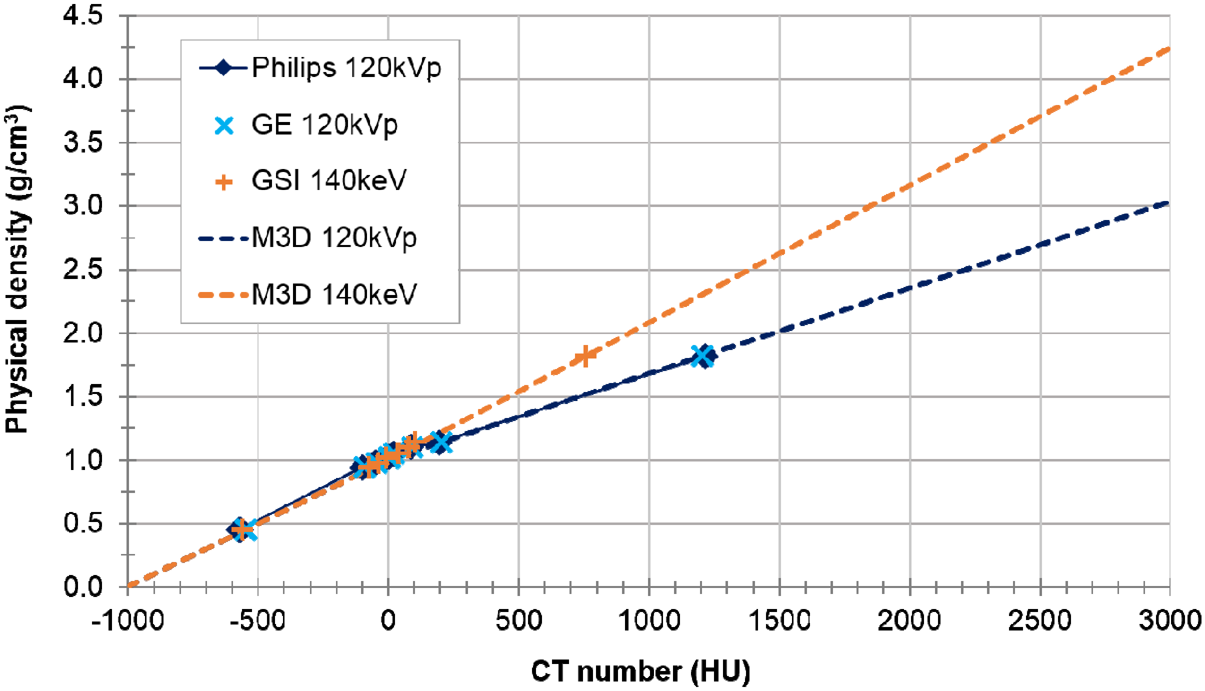


Figure 5.2: CT number-to-density conversion curves for 120kVp imaging (“M3D 120kVp”) and monochromatic 140keV GSI imaging with and without MARS (“M3D 140keV”). The data points labeled “Philips 120kVp,” “GE 120kVp,” and “GSI 140keV” show data obtained from scans of the Gammex tissue characterization phantom that were used to create the CT number-to-density conversion curves.

Metal artifact reduction study

In order to quantify the dosimetric impact of using various CT metal artifact reduction methods for photon dose calculations, we imaged the phantom described in Section 2.A using baseline, uncorrected 120kVp imaging, and O-MAR, GSI, and MARs artifact reduction imaging. All protocols for both the titanium and Cerrobend configurations of the phantom are listed in Table 5.1. Each phantom was scanned using both a Philips Brilliance (Cleveland, OH) and a GE Discovery CT750 HD (Milwaukee, WI) scanner. For the Philips scanner, the phantom was scanned using a 120kVp protocol and reconstructed with and without the O-MAR algorithm. For the GE scanner, the phantom was scanned using polyenergetic imaging (120kVp) and with its dual-energy mode, which allowed for monoenergetic image reconstruction. For the GSI dual-energy scan, images were reconstructed at 140keV, the highest energy available, with and without the MARs algorithm. The 140keV energy level was chosen based on previous work that indicated that 140keV was more successful than lower energies at reducing metal artifacts.[68] In order to draw a fair comparison between the two vendors, we matched the CT protocols based on various acquisition and reconstruction parameters. All reconstructed images were 12-bit depth images. In summary, phantom images were obtained using baseline, uncorrected imaging methods ("Philips 120kVp" and "GE 120kVp") and the artifact reduction methods under investigation ("O-MAR," "GSI 140keV," and "MARs 140keV").

These phantom images sets were then used to create treatments plans to match the measurement conditions described in section 2.B (6-MV photons, 5 cm x 5 cm field size, 100-cm source to surface distance) using the Pinnacle^3 v9.0 treatment planning system. All DICOM files associated with the plan, including CT images, RTPLAN, RTSTRUCT, and RTDOSE files, were then exported from Pinnacle to the M3D system to perform dose calculations with and without the metal kernel implementation, using a 1.5 mm uniform dose grid. The M3D-calculated dose was then compared to the measured dose, and a mean % error was reported for both the upstream region (from d_{max} to the proximal interface) and the downstream region (from the distal interface to the deepest point of measurement, approximately 5 cm beyond the metal implant).

Table 5.1: Scan protocols for phantom scans using the Philips Brilliance and the GE Discovery CT750 HD CT scanners. All protocols are helical scans.

Phantom Insert	Protocol	Pitch	mA	Tube rotation time (s)	Filter	Slice thickness (cm)	DFOV (cm)	Recon kernel	CTDIvol (mGy)
Titanium	Philips 120kVp	0.938	231	0.75	B	3.0	50	B	12.3
	GE 120kVp	0.984	265	0.6	Medium	2.5	50	Standard	13.2
	GE GSI preset #41	0.984	360	0.6	Medium	2.5	50	Standard	13.1
Cerrobend	Philips 120kVp	0.688	206	0.75	B	3.0	50	B	26.7
	GE 120kVp	0.516	280	0.6	Medium	2.5	50	Standard	26.8
	GE GSI preset #32	0.516	375	0.6	Medium	2.5	50	Standard	26.4

Metal kernel study

In order to study the impact of implementing metal kernels for photon dose calculations without the confounding effects of the imaging artifacts, we chose to perform dose calculations using metal kernels in an “ideal” phantom geometry. For both phantom configurations (titanium and Cerrobend), this ideal geometry was created using CT images acquired as described in section 2.D but modified with density overrides for the metal insert and the streak artifacts in the high-impact polyethylene portion of the phantom, based on the known dimensions of the phantom geometry and the known densities of the materials. The dose was then calculated using the M3D collapsed cone algorithm using both the traditional implementation of density scaling of water kernels and the novel implementation using metal kernels. For comparison, the dose for these ideal treatment plans was also calculated using the Pinnacle CCC and the Eclipse AAA algorithms (Varian Medical Systems, Palo Alto, CA). Furthermore, to investigate dose grid size dependence, we also performed M3D metal kernel dose calculations with a variety of uniform dose grid sizes (1.25 mm, 1.5 mm, and 3.0 mm). All dose calculations were compared against the dose measured with film, and the mean % error within 1 cm of each interface was reported. Based on previous work,[65] the dosimetric impact of metal kernels is a fairly local effect that does not extend further than approximately 1 cm from the interface for 6-MV photons.

Artifact reduction methods and metal kernels combined study

To investigate the dose calculation accuracy that can be achieved when metal artifact reduction methods are combined with the use of metal kernels in the dose calculation algorithm, we performed dose calculations using the M3D system for the possible combinations of these two strategies: baseline calculation (no artifact reduction or metal kernels), artifact reduction calculation, metal kernel calculation, and artifact reduction with metal kernel calculation. These dose calculations were performed using the M3D system with a 1.5-mm uniform dose grid and the CT images of the phantom acquired as described in section 2.D. The mean % error between calculated and measured dose is reported for the upstream region (from d_{max} to the proximal interface) and the downstream region (from the distal interface to the deepest point of measurement, approximately 5 cm beyond the metal implant).

5.3 Results

Metal artifact reduction study

As described in section 2.D, the dosimetric impact of CT metal artifact reduction methods was investigated by acquiring images of the phantom with titanium and Cerrobend inserts using both uncorrected, baseline imaging techniques (Philips 120kVp and GE 120kVp) and all three metal artifact reduction methods (O-MAR, GSI 140keV, and MARs 140keV); we then performed depth dose calculations using these image sets (Figure 5.3). Table 5.2 lists the mean absolute % errors between the calculated and measured dose in the upstream and downstream regions. On the baseline scans, substantial errors were observed. The average error upstream was only 1-2% for titanium and 7-10% for Cerrobend, but at the interface the dose error exceeded 20% for titanium and 50% for Cerrobend. Large systematic errors were seen in the downstream region, on average more than 10% for titanium and more than 30% for Cerrobend.

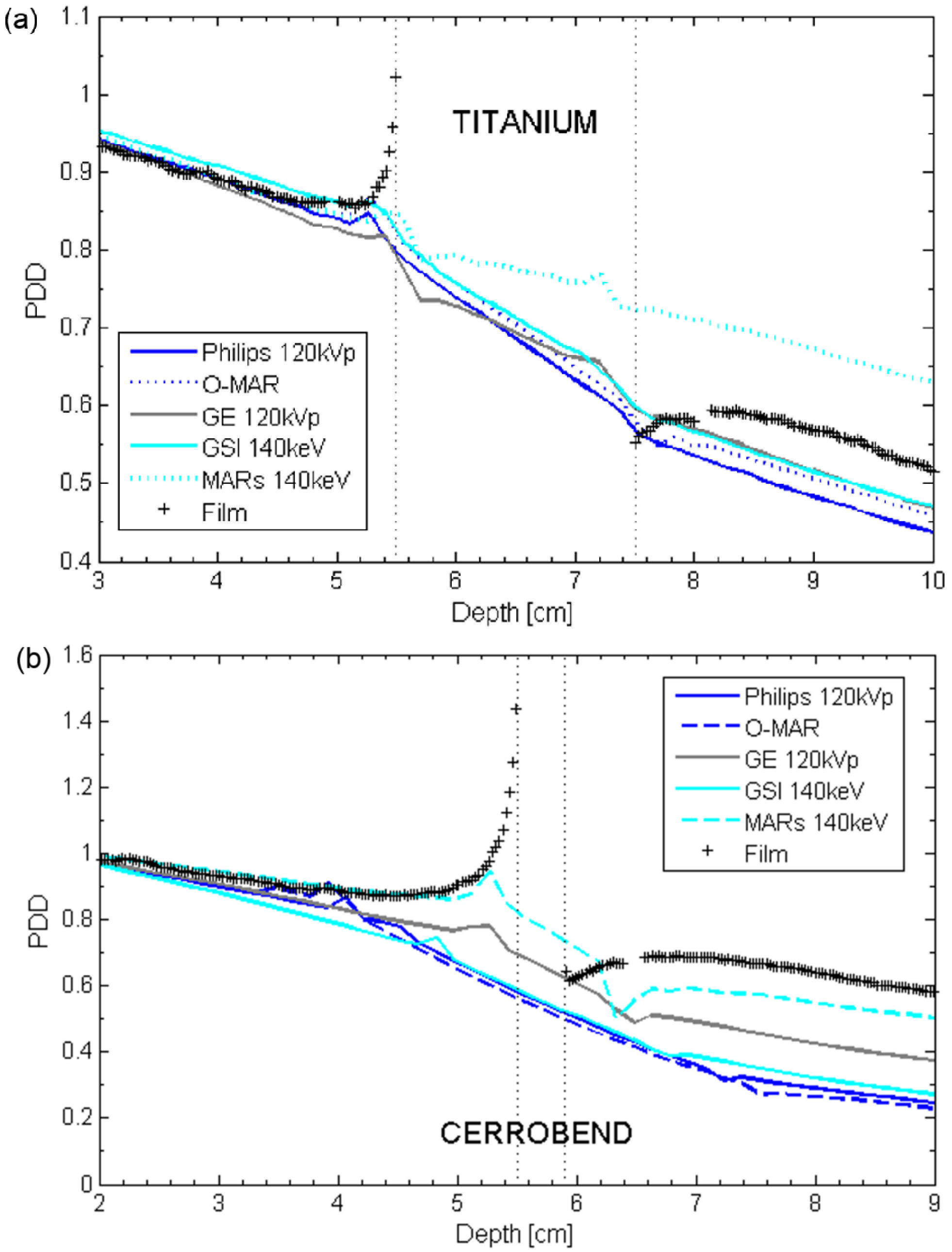


Figure 5.3: Percent depth dose (PDD) curves calculated for (a) titanium and (b) Cerrobend using baseline imaging techniques (Philips 120kVp and GE 120kVp) and metal artifact reduction methods (O-MAR, GSI 140keV, and MARs 140keV). The dose measured with film is also shown.

Table 5.2: The mean absolute % error between calculated and film-measured dose for the titanium and Cerrobend phantoms. Dose calculations were performed using various baseline imaging methods and metal artifact reduction methods. The upstream region extends from d_{max} to the proximal interface, while the downstream region extends from the distal interface to 5 cm beyond the metal implant.

Imaging technique	Titanium: Mean Absolute % Error		Cerrobend: Mean Absolute % Error	
	Upstream	Downstream	Upstream	Downstream
Philips 120kVp	1.4	15.0	10.1	53.7
O-MAR	1.1	11.1	10.3	56.4
GE 120kVp	1.9	9.0	7.3	33.0
GSI 140keV	1.9	8.7	12.0	49.6
MARs 140keV	1.4	21.8	2.5	14.7
MARs 140keV*	1.3	2.0	--	--

*Density override was performed for the low-density pocket caused by the MARs algorithm.

Application of the artifact reduction methods often provided some benefit, but sometimes produced substantially erroneous results. For the titanium phantom, O-MAR was successful in reducing the error in both the upstream region (from 1.4% to 1.1%) and the downstream region (from 15.0% to 11.1%) in comparison to uncorrected, baseline imaging. GSI 140keV imaging had no effect on the overall error in the upstream region but was able to reduce the error in the downstream region slightly (from 9.0% to 8.7%). The MARs algorithm was successful in reducing error in the upstream region; however, it greatly increased the error in the downstream region in comparison to baseline imaging (from 9.0% to 21.8%). This large increase in the error arose from the artificial creation of a low-density pocket within the titanium implant by the MARs algorithm (Figure 5.4b). If no manual intervention was performed to override the incorrect densities assigned to this pocket (Figure 5.4c), the attenuation caused by the titanium implant was highly underestimated, resulting in a substantial overestimation of dose downstream of the metal. However, when the correct density was assigned to the low-density pocket, the dose calculation accuracy in the downstream region was greatly improved, yielding a mean absolute error of 2.0%. This accuracy could be achieved because the algorithm was otherwise successful at reducing streak artifacts and at accurately representing the external dimensions of the titanium implant. The streak artifacts in the uncorrected, baseline images of the titanium implant result in an overestimation of the density along the central axis, with this overestimation getting worse as you get closer to the metal implant, while the MARs algorithm corrected this behavior (Figure 5.4c). Figure 5.5 illustrates how manually overriding the density of the pocket caused by MARs can greatly improve dose calculation accuracy.

The 4mm-thick Cerrobend implant was a much more challenging case for dose calculation in general. Most of the imaging methods overestimated the attenuation caused by the Cerrobend implant as seen in Figure 5.3. This overestimation of attenuation was caused in part by the metal streak artifacts, which artificially inflated density assignment values near the metal, but the dominant factor was actually the misrepresentation of the metal implant's size in the images. Most of the imaging methods, including artifact reduction methods, severely overestimated the thickness of the Cerrobend implant and thus overestimated the attenuation caused by the metal. The only artifact reduction method that was able to reduce dose calculation errors was the MARs algorithm. In comparison to GE baseline imaging, which overestimated the thickness of the Cerrobend (Figure 5.4d), the MARs algorithm reduced the severe artifacts caused by the high-density metal (Figure 5.4e) and more accurately represented the thickness of the implant (Figure 5.4f). Accordingly, in the upstream region, MARs reduced the mean absolute % error from 7.3% to 2.5% in comparison to uncorrected, baseline imaging, while in the downstream region, MARs reduced the error from 33.0% to 14.7%. MARs did not artificially create a low-density pocket in the Cerrobend insert, probably because the insert was much thinner than the titanium insert.

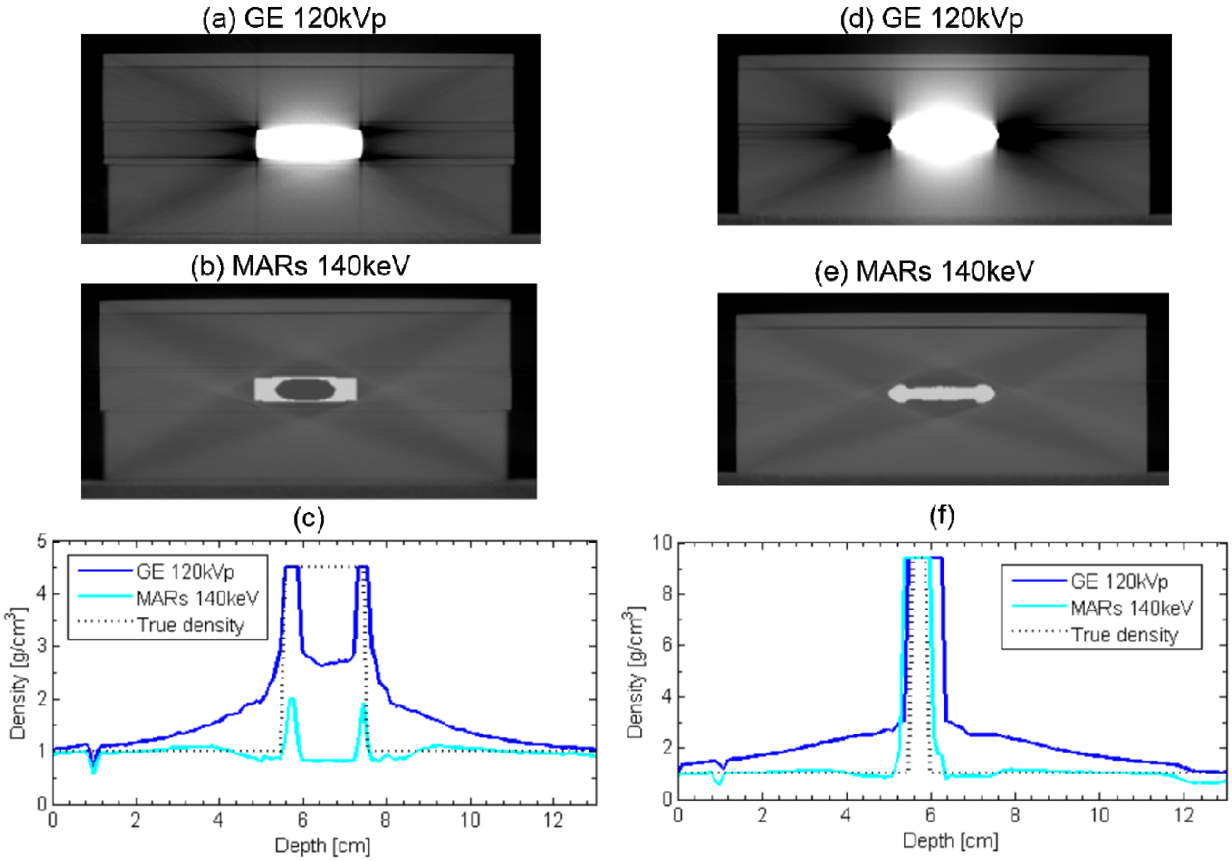


Figure 5.4: Panels (a) and (d) show 120kVp images of the titanium and Cerrobend phantom configurations, respectively, while panels (b) and (e) show the same image slice using MARs 140keV imaging, illustrating the artificial introduction of a low-density pocket within the titanium in (b). Panels (c) and (f) show the density assignment along the central axis using both imaging methods, along with the true density.

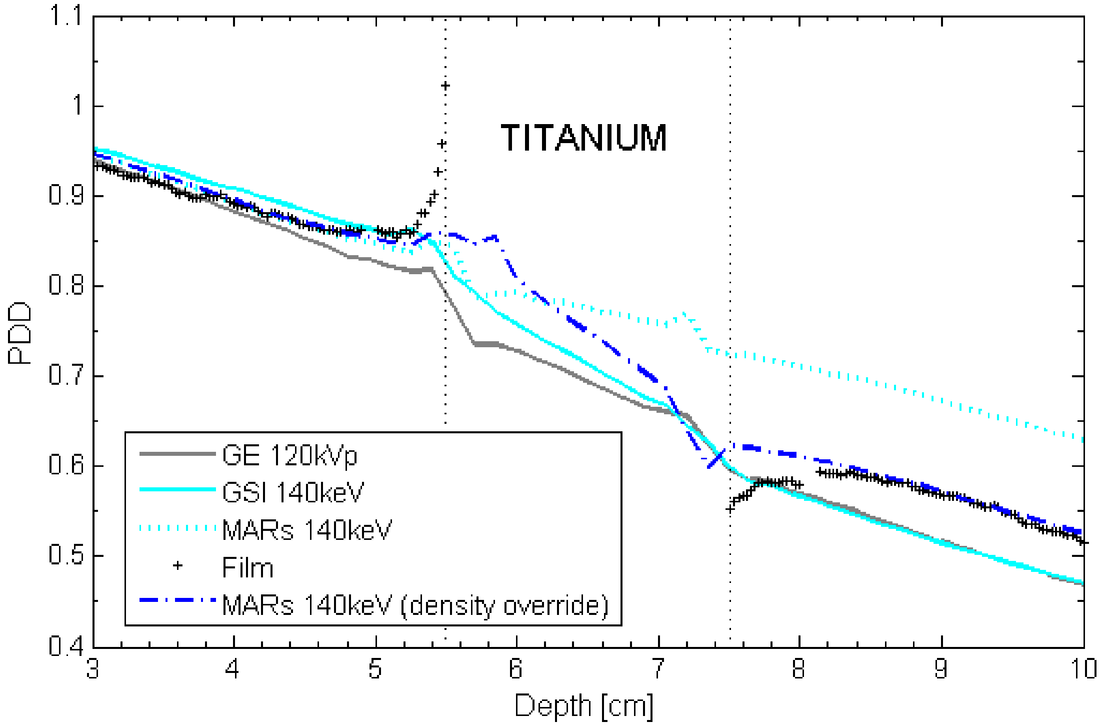


Figure 5.5: Percent depth dose (PDD) curves for the titanium insert illustrating the effect of overriding the incorrect low-density values assigned to the inner portion of the implant caused by the MARs algorithm. The dose measured with film, as well as the dose calculated using 120kVp imaging, GSI 140keV imaging, and MARs 140keV imaging without any manual intervention are shown for comparison.

Metal kernel study

As described in section 2.E, the dosimetric impact of implementing metal kernels without the confounding effects of imaging artifacts was investigated using a virtual (i.e., “ideal”) phantom geometry. The percent depth dose curves for M3D dose calculations with and without metal kernels, along with the dose calculated with Pinnacle’s CCC and Eclipse’s AAA algorithms, are shown in Figure 5.6. The mean absolute % errors in the 1-cm region upstream of the proximal interface and the 1-cm region downstream of the distal interface for titanium and Cerrobend are listed in Table 5.3. These data reveal that metal kernels generally improved dose calculation accuracy in the upstream region. For instance, for Cerrobend, metal kernels reduced the upstream error from 9.2% to 8.2% (1.25-mm uniform dose grid). However, metal kernels did not universally improve dose calculation accuracy, and in the downstream region, they generally worsened it. For titanium, implementation of metal kernels increased the error from 2.4% to 3.7% in the downstream region. Also, from Figure 5.6, it is clear that the different commercial dose calculation algorithms showed markedly different behaviors at the sharp metal/tissue interfaces. For titanium, M3D modeled the backscatter peak at the proximal interface better than the Pinnacle or the Eclipse algorithms (Figure 5.6a). For the distal interface of both metals, M3D exhibited a sharp dose drop-off in the re-buildup region, while the Pinnacle and Eclipse algorithms showed a more gradual transition. Furthermore, even though all three algorithms resulted in similar percent depth dose values past the re-buildup region (>1.5 cm past the distal interface), each algorithm exhibited distinct behavior for dose deposition inside the metal itself.

In addition to investigating whether metal kernels improved dose calculation accuracy, we also investigated whether the accuracy of metal kernel dose calculations was dependent on dose grid size. Depth dose curves calculated using M3D with and without metal kernels for various dose grid sizes are shown in Figure 5.7, and the mean absolute % error for the upstream and downstream 1-cm regions are listed in Table 5.3. Metal kernels generally yielded more accurate calculations as the dose grid size decreased. For instance, for a 3.0-mm dose grid size, metal kernels worsened dose calculation accuracy downstream of Cerrobend (error increased from 3.7% to 4.3% relative to the dose calculation with water kernels only), while for a 1.25-mm dose grid, metal kernels improved the accuracy for the same calculation geometry (error decreased from 2.6% to 2.0%).

Table 5.3: The mean absolute % error between calculated and film-measured dose for dose calculations performed with M3D with and without metal kernels (MK), Pinnacle CCC, and Eclipse AAA. Dose calculations were performed for virtual ("ideal") phantom geometries. The % error is reported for the region 1 cm upstream of the proximal interface and the region 1 cm downstream of the distal interface.

Dose calculation method	Dose grid size (mm)	Titanium: Mean Absolute % Error		Cerrobend: Mean Absolute % Error	
		Upstream 1 cm	Downstream 1 cm	Upstream 1 cm	Downstream 1 cm
M3D (no MK / MK)	1.25	1.5 / 1.2	2.4 / 3.7	9.2 / 8.2	2.6 / 2.0
M3D (no MK / MK)	1.5	1.4 / 1.8	2.2 / 6.1	9.0 / 7.5	2.8 / 6.6
M3D (no MK / MK)	3.0	1.7 / 1.8	4.5 / 7.9	9.4 / 8.2	3.7 / 4.3
Pinnacle CCC	1.5 x 1.5 x 1.0	2.6	1.6	9.2	9.2
Eclipse AAA	2.5	2.2	5.0	9.8	6.3

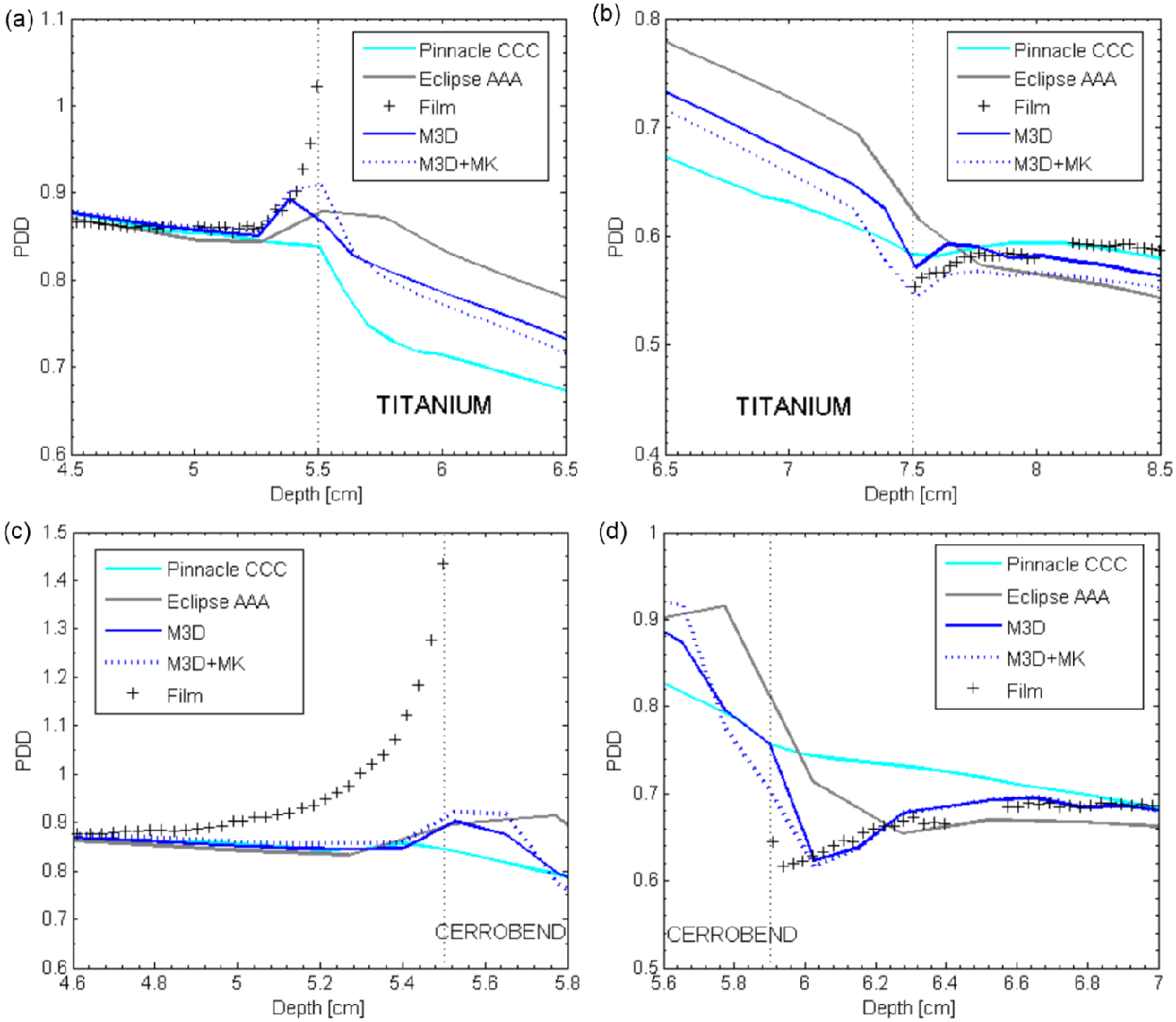


Figure 5.6: Percent depth dose (PDD) curves for dose calculations performed using M3D with (“M3D+MK”) and without (“M3D”) metal kernels, Pinnacle CCC, and Eclipse AAA at (a) the proximal interface of the titanium insert, (b) the distal interface for titanium, (c) the proximal interface for Cerrobend, and (d) the distal interface for Cerrobend. The dose measured with film is also shown. Dose calculations for M3D were performed for a uniform 1.25-mm dose grid.

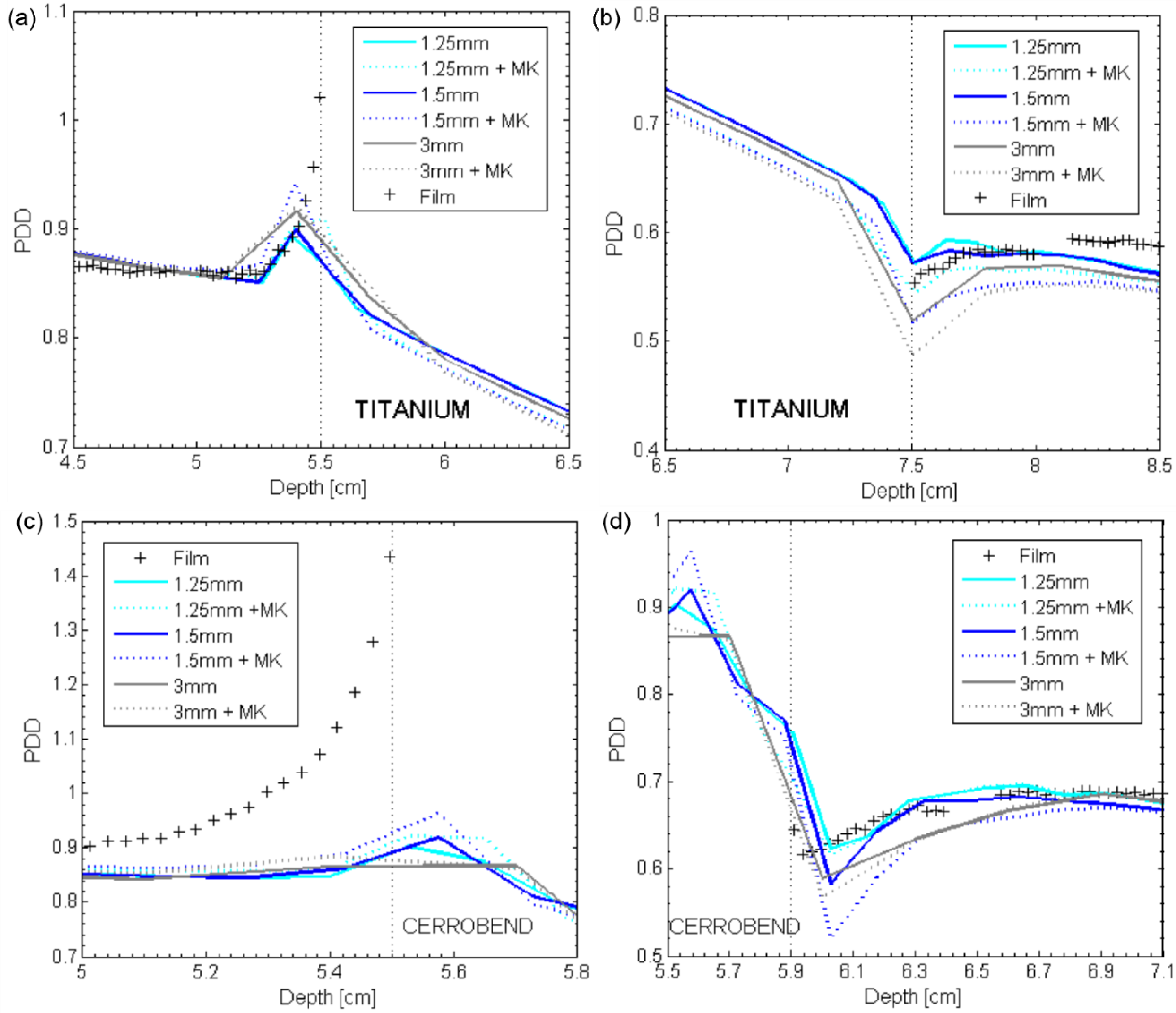


Figure 5.7: Percent depth dose (PDD) curves for dose calculations performed using M3D with metal kernels ("+MK") and without metal kernels using various dose grid sizes for (a) the proximal interface of titanium, (b) the distal interface of titanium, (c) the proximal interface of Cerrobend, and (d) the distal interface of Cerrobend. The dose measured with film is also shown.

Artifact reduction methods and metal kernels combined study

The results of our study combining artifact reduction methods and metal kernels are summarized in Table 5.4. When artifact reduction methods and metal kernels were used in conjunction, we found that the resulting accuracy that could be achieved was essentially additive. If the artifact reduction method and metal kernels were each successful at reducing dose calculation errors individually, then using them in combination resulted in the best accuracy overall. An example of this can be seen for the region upstream of titanium when using

O-MAR. Using O-MAR reduced the error from 1.4% to 1.1% compared to uncorrected imaging, and using both O-MAR and metal kernels further reduced the error to 0.9%. However, in cases where metal kernels were unsuccessful at improving accuracy (i.e., in the downstream region), using an artifact reduction method with metal kernels did not give the best accuracy overall. Unfortunately, results were mixed, and neither error-reduction strategy was universally beneficial. For titanium, application of a metal artifact reduction strategy was generally beneficial both upstream and downstream, while metal kernels were generally beneficial only upstream of this metal. However, the MARs algorithm could produce disastrously erroneous results if the artificially created low-density pocket was not identified and corrected. For the Cerrobend case, conversely, only the application of MARs made any substantial improvement in calculation accuracy.

Table 5.4: The mean absolute % error between calculated and film-measured dose for the four possible combinations of the two error reduction strategies (CT metal artifact reduction [AR] methods and metal kernels [MK]). The mean % error is reported for the upstream region (extending from d_{max} to the proximal interface) and the downstream region (extending from the distal interface to 5 cm past the metal). The baseline imaging method was Philips 120kVp for O-MAR, while the baseline method was GE 120kVp for GSI 140keV and MARs 140keV.

Metal	AR method	Upstream region				Downstream region			
		Baseline (120kVp)	AR	Baseline +MK	AR +MK	Baseline (120kVp)	AR	Baseline +MK	AR +MK
Titanium	O-MAR	1.4	1.1	1.2	0.9	15.0	11.1	16.4	12.7
	GSI 140keV	1.9	1.9	1.8	1.8	9.0	8.7	9.3	9.1
	MARs 140keV*	1.9	1.3	1.8	1.3	9.0	2.0	9.3	1.1
Cerrobend	O-MAR	10.1	10.3	10.0	10.4	53.7	56.4	54.7	57.6
	GSI 140keV	7.3	12.0	7.1	11.9	33.0	49.6	33.7	50.4
	MARs 140keV	7.3	2.5	7.1	2.4	33.0	14.7	33.7	15.9

*Density override was performed for the low-density pocket caused by the MARs algorithm.

5.4 Discussion

In this study, we investigated the dosimetric impact of several commercial CT metal artifact reduction methods and the impact of implementing metal energy deposition kernels in a commercial collapsed cone C/S algorithm. Both strategies for reducing dose calculation errors for patients with metal implants were evaluated using a simple slab phantom geometry containing inserts for two metals representative of those frequently encountered in radiotherapy. Our results indicate that although both strategies can improve dose calculation accuracy in some cases, neither is globally effective, and users should be aware of the scenarios in which these strategies may worsen dose calculation accuracy.

In general, we found that the various metal artifact reduction methods were more successful for the titanium implant geometry than for Cerrobend. The fact that the artifact reduction methods were more successful for titanium is unsurprising since Cerrobend is a very dense metal with a high effective atomic number (76.8).[10] Due to the highly attenuating properties of Cerrobend, photon starvation was a greater issue for this metal, resulting in very limited information for projections involving the metal implant. It should be noted that although the thickness of the Cerrobend implant used in this study (4 mm) is realistic for dental restorations, the lateral dimensions of the implant were larger than would be encountered clinically for a single filling but rather would more closely mimic a row of fillings. Though the Cerrobend geometry in this study was chosen somewhat for its convenience with measurements, it may have been an overly challenging case for CT metal artifact reduction. Thus, a limitation of this study is that the Cerrobend implant results may only be partly applicable to the clinical case of dental fillings. However, additional investigations are underway to evaluate the dosimetric benefit of using these artifact reduction methods for a more clinically realistic geometry (dental amalgam in an anthropomorphic phantom).

One interesting result from this study is that the accuracy of dose calculations for beams traversing a metal implant is highly dependent on how accurately the size and dimensions of the metal implant are represented in the CT images. Not only can imaging artifacts cause inaccurate density assignment near the metal implants, but also, for very dense metals such as Cerrobend, the CT numbers in the image can saturate, not only inside the metal itself but also in nearby pixels. This blooming or spreading of saturated CT numbers causes the physical dimensions of the implant to be overestimated. If there is no prior knowledge of the dimensions of the implant that can be used to perform manual density overrides, and if extended bit-depth CT images are not available, this overestimation of metal size can lead to large dose calculation errors. The dosimetric impact of metal size accuracy was especially evident from our Cerrobend data, in which uncorrected CT images resulted in systematic errors greater than 30% in the downstream region. Furthermore, the success of the metal artifact reduction methods was highly dependent on if the method was able to reduce imaging artifacts while still accurately representing the size and dimensions of metal implants. In fact, even if a method was able to reduce imaging artifacts, those artifacts may have been only a tiny component to the dose calculation inaccuracy. Proper representation of the size of the metal implant was highly variable between approaches and made a substantial difference to dose calculation accuracy. In the most dramatic instance, we observed that the MARs algorithm caused a low-density pocket in the center of the titanium implant in this study (Figure 5.4b). This behavior of the MARs algorithm to create low-density pockets[68] and distort the size/shape of metal implants has been documented previously.[43-45] The clinical implications of this behavior may be profound if left uncorrected: a resulting overestimation of dose downstream, for instance, could lead to unacceptably high dose to organs at risk in the beam path. In our study, use of the MARs algorithm for titanium resulted in higher dose errors than use of uncorrected CT images, when the low-density pocket was left uncorrected. Consequently, if little information is known about the composition and construction of the metal implant, it may be better to use no metal artifact reduction strategy than to risk these large dose errors associated with MARs metal distortion. In contrast, the MARs algorithm resulted in substantial improvements in metal thickness accuracy and dose calculation accuracy for our Cerrobend insert, indicating that the success of MARs may be dependent on the type of metal and the size of the implant, a feature that may be unattractive given the wide variety of implants that can be encountered in radiation oncology.

In addition to CT metal artifact reduction methods, we also investigated the dosimetric impact of implementing metal kernels in a commercial collapsed cone C/S dose calculation algorithm. The dosimetric impact of metal kernels, in comparison to traditional density scaling of water kernels, is a local effect; the dose differences between dose calculations with and without metal kernels only extend up to 1 cm from the interface. Unfortunately, metal kernels did not universally improve dose calculation accuracy. In comparison to dose calculations using water kernels only, metal kernels generally improved the accuracy at the proximal interface but worsened accuracy at the distal interface. Metal kernels exhibited improved backscatter modeling, which may be beneficial in predicting and preventing oral mucositis, an adverse side effect of head and neck radiation therapy that can occur due to the local dose enhancement caused by dental restorations.[6] However, the decreased accuracy directly downstream of metals may be a deterrent for using metal kernels if there is a target or organ at risk directly downstream of a metal implant. Because our study only investigated a simple geometry with a single photon beam, it is not clear what the clinical impact of metal kernels will be for more complex treatment plans. It is possible that for IMRT treatment plans, with many beam angles, the dosimetric advantages (upstream) and disadvantages (downstream) metal kernels will cancel each other out somewhat. Investigation of the use of metal kernels for clinical treatment plans is the topic of future research.

It is also important to note that we observed that several commercial algorithms (M3D, Pinnacle CCC, and Eclipse AAA) handled the sharp metal/tissue interfaces differently. It is possible that markedly different results would be obtained if metal kernels were implemented into the Pinnacle CCC algorithm for instance as opposed to the M3D algorithm. With the metal kernels, the M3D-calculated dose is decreased compared to the dose values calculated with water kernels, resulting in percent depth dose values that were too low downstream of the distal interface. However, if metal kernels were implemented in the Pinnacle CCC algorithm, a decrease in dose due to metal kernels would improve the accuracy in this region. It is clear that subtle differences in the implementation of the collapsed cone C/S method between the Pinnacle and M3D algorithms can have a major impact on the calculated dose near sharp metal/tissue interfaces and that these implementation differences can have a larger dosimetric impact than metal kernels.

We also performed dose calculations with various dose grid sizes to see if the accuracy of metal kernel dose calculations had any dose grid size dependence. Intuitively, one would expect that using smaller calculation voxels would better model interface effects. In agreement with expectations, we found that smaller calculation voxels resulted in better accuracy for metal kernels. Unfortunately, we were not able to decrease the dose grid size below 1.25 mm due to software and memory limitations. However, our data indicate that if metal kernels are to be implemented in a commercial treatment planning system, a smaller dose grid size would be preferable in terms of accuracy, though at the expense of calculation speed.[27]

A limitation of our study is that all dose calculations were performed in a geometric phantom with a single incident photon beam; thus, our analysis does not include all of the complexities of modern patient geometries and treatments. Therefore, we have already initiated work to investigate the impact of metal artifact reduction methods and metal kernels for clinical scenarios (the case of a patient with dental fillings receiving head and neck radiation therapy and the case of a patient with spinal hardware receiving stereotactic spine radiosurgery). Another possible area of future work is the investigation of metal artifact reduction methods and use of dual-energy CT for proton therapy dose calculations. Other investigators have shown that proton therapy dose calculations are more sensitive to streak artifacts than photons because the artifacts cause proton range errors that can lead to overdosing of normal structures distal to the target.[72, 73]

5.5 Conclusions

In this study, we found that in comparison to dose calculations with water kernels only, use of metal kernels resulted in better modeling of the increased backscatter dose at the proximal metal interface but decreased dose calculation accuracy directly downstream of the metal. The success of the metal kernels for improving dose calculation accuracy was also found to be dependent on the resolution of the dose calculation grid, with finer resolution generally yielding better accuracy. For the commercial CT metal artifact reduction methods investigated, it was found that the methods were generally more successful for titanium than Cerrobend. Interestingly, we also found that how the artifact reduction methods affect metal size may be more important than how successful they are at reducing the streak artifacts emanating from the metal. For the titanium implant, the MARs algorithm distorted the metal and resulted in a hypointense region in the center of the implant and a substantial decrease in calculation accuracy. However, for the Cerrobend implant, the MARs algorithm was successful at reducing the severe streak artifacts and more accurately represented the thickness of the Cerrobend than uncorrected imaging, resulting in substantial gains in accuracy. Because of this inconsistent behavior, the MARs algorithm should be used cautiously for treatment plans in which beams traverse a metal implant and little is known about the composition of the implant. In summary, both CT metal artifact reduction methods and use of metal kernels in the dose calculation algorithm resulted in improvements in dose calculation accuracy in some cases, but neither strategy was globally effective. If these strategies are to be implemented clinically, users should be aware of scenarios in which these strategies may worsen calculation accuracy.

Chapter 6: Strategies for reducing dose calculation errors for radiation therapy patients with metal implants near the treatment site

6.1 Introduction

External beam radiation therapy is commonly delivered to patients who have a metal implant near the site of the targeted tumor. This includes the 1-4% of all radiation therapy patients who have a prosthetic device that can affect their treatment,[1] as well as the large number of head and neck cancer patients who have dental restorations. These metal implants pose challenges in the treatment planning process and can degrade the accuracy of the resulting treatment plans. Dose calculation errors associated with the metal implants arise from limitations of computed tomography (CT) imaging, as well as from the inability of modern dose calculation algorithms to model complicated metal/tissue interface effects.

Metal implants cause artifacts that degrade the quality of CT images used for treatment planning. Not only do these artifacts make it difficult to visualize the diseased area, they also cause errors in the CT numbers that are used for density assignment in heterogeneous dose calculations.[15, 16] These density assignment errors can lead directly to dose calculation errors. Recently, several commercial metal artifact reduction solutions have become available, and commercial algorithms are becoming increasingly popular in both diagnostic imaging and radiation oncology. The Philips O-MAR algorithm (Philips Healthcare, Cleveland, OH) is an iterative projection modification solution, which identifies projection data affected by the metal implant and replaces the corrupted data using nearby uncorrupted data.[36] Unlike the O-MAR algorithm, which is a software-only approach using polyenergetic CT data, Gemstone Spectral Imaging (GSI) uses dual-energy CT data. The Discovery CT750 HD system (GE Healthcare, Milwaukee, WI) uses a single X-ray source that rapidly switches between 80 and 140kVp to acquire alternating low and high energy projection data.[38] These dual-energy projection data are used to generate virtual monochromatic images (with energy levels between 40keV and

140keV) that depict how an object would look if it were imaged using a monoenergetic X-ray source.[42] In comparison to polyenergetic images, these virtual monochromatic images exhibit reduced beam hardening artifacts, especially at the higher reconstruction energy levels.[38, 68] Additionally, GE has also developed a metal artifact reduction software (MARs) specifically for use with GSI monochromatic imaging. It combines the benefits of virtual monochromatic imaging with the benefits of a dedicated artifact reduction algorithm. The MARs algorithm has shown benefits for reducing artifacts caused by an orthopedic prosthesis,[43, 68] spinal screws,[44] and fiducial markers.[45]

In addition to limitations of CT imaging, dose calculation errors are also caused by the inability of modern dose calculation algorithms to accurately model complicated interface effects. The convolution/superposition (C/S) dose calculation method, which convolves the energy released via primary photon interactions with an energy deposition kernel, is the current standard of care for photon dose calculations. The energy released in each voxel is described by the TERMA (total energy released per unit mass), while the spatial distribution of energy deposition via secondary electrons and scattered photons is described by the energy deposition kernel. While C/S methods can accurately calculate the attenuation caused by a metal implant in ideal cases (in which the density and physical dimensions of the implant are well defined), C/S methods result in dose calculation errors near the metal/tissue interfaces, underestimating the backscatter dose enhancement at the proximal interface and overestimating the dose directly downstream of the implant.[1] Although these errors are typically confined to the local region near metal/tissue interfaces (within a few centimeters of the interfaces), local errors can be as high as 30%, depending on the photon energy and the type of metal.[10, 11] One source of this dose calculation error is that for commercial C/S algorithms, the energy deposition kernels are based on photon interactions and scatter in water, and these water-based kernels are simply scaled in dimension based on the local density to perform heterogeneous dose calculations (density scaling approximation). Though this approximation is reasonable for tissue, it is less accurate for materials that differ greatly from water in their atomic composition.[33] Recently, our group Huang, et al.[65] generated kernels for several metals and highlighted that the shape of metal kernels differed from that of water kernels, indicating that density scaling of water kernels is inadequate to describe photon interactions in metals. We also found, promisingly, that implementing these metal kernels in a commercial algorithm resulted in increased backscatter dose and decreased dose directly downstream of metals. Therefore, metal kernels have the potential to correct the calculation errors at metal/tissue interfaces that have been observed with conventional C/S algorithms.

Although we have previously investigated metal kernels,[65] as well as the success of commercial CT artifact reduction methods for several common implants,[68] the magnitude of the improvement in dose calculation accuracy due to these strategies is not known for clinically realistic implants and treatments. Thus, the purpose of this study was to investigate these error-reduction strategies using two clinical cases. In the dental case, we investigated the benefits of these two error-reduction strategies for an oral cavity treatment with an anthropomorphic phantom with amalgam fillings. In the spine case, we investigated calculation accuracy for a stereotactic spine radiosurgery treatment in the presence of titanium spinal rods. For both clinical cases, beam arrangements traversing the metal implant are difficult to avoid, and dose calculation accuracy is especially important as organs at risk and tumors can be very close to the metal implant.

6.2 Methods and Materials

Anthropomorphic phantoms

To mimic a patient with spinal hardware, a thoracic anthropomorphic phantom was used. This thoracic phantom was designed by the Imaging and Radiation Oncology Core (IROC) Houston Quality Assurance Center, formerly known as the Radiological Physics Center, and contains structures representing the heart and lungs.[63] For this study, this spine insert of this phantom, which contains structures representing the spinal cord, bone, and esophagus, was replaced with a high impact polystyrene insert that could accommodate two titanium rods (9.5 mm in diameter) to mimic spinal fixation rods, as well as four Exradin A1SL 0.057cc ion chambers (Standard Imaging, Middleton, WI) and an axial film plane that intersected the titanium rods. For some of our measurements, the titanium rods were replaced with high impact polystyrene rods to investigate the level of accuracy that could be achieved with no metal implants (no-metal case).

To mimic the case of a head and neck cancer patient with dental fillings, we custom-built a phantom; this phantom contained two halves (mimicking an upper and lower jaw) composed of high impact polystyrene. Each half of the phantom contained tooth structures made of Gammex 450 cortical bone substitute (Middleton, WI). One set of teeth contained cortical bone material only and was used to quantify the level of dose calculation accuracy that could be achieved without metal, while a second set of teeth was modified to contain Dispersalloy® dental amalgam (Dentsply, Milford, DE). For the set of teeth modified to hold dental amalgam, 12 fillings were simulated, 6 in the upper set of teeth and 6 in the lower set of teeth. The locations and dimensions of the fillings were selected by a dental oncologist to be clinically realistic (Figure 6.1). As with the spine phantom, the dental phantom was designed to accommodate four Exradin A1SL ion chambers at various locations in the oral cavity, as well as an axial film plane between the lower and upper jaws of the phantom.

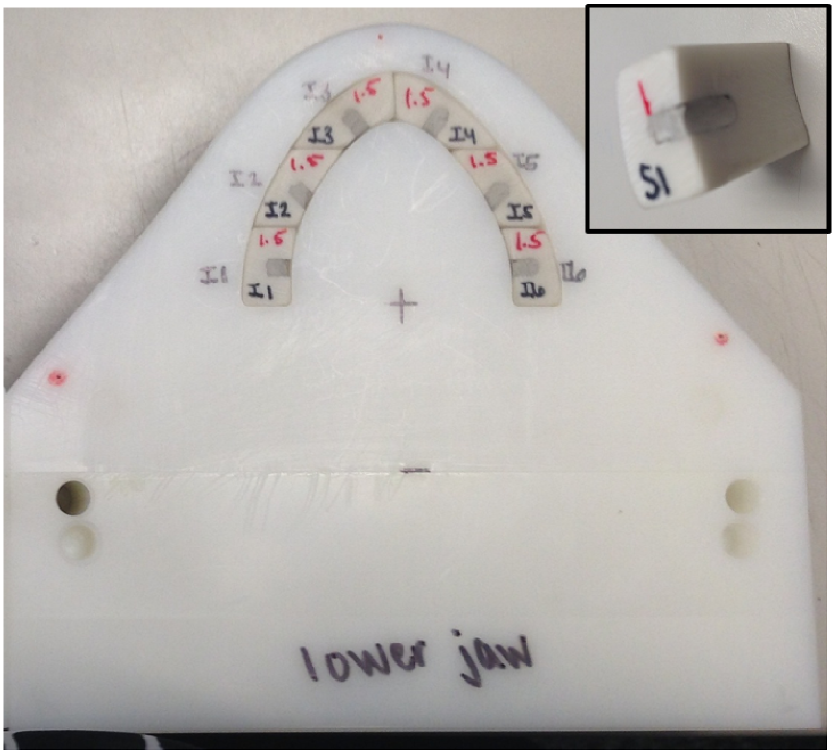


Figure 6.1: A photograph of half of the dental phantom illustrating the cortical bone tooth structures that have been modified to hold amalgam dental fillings. The inset shows an individual tooth structure with dental amalgam.

Imaging

To study the dosimetric impact of using CT metal artifact reduction methods, the spine and dental phantoms (with and without the metal implants in place) were imaged using baseline, uncorrected 120kVp imaging on both a Philips Brilliance and a GE Discovery CT750 HD scanner. The phantoms with metal in place were also imaged using artifact reduction methods. For the Philips scanner, the phantom was scanned using a 120kVp protocol (designated in this report as “Philips 120kVp”) and reconstructed with and without the O-MAR algorithm. For the GE scanner, the phantom was scanned using polyenergetic imaging (baseline imaging, “GE 120kVp”) and also with a dual-energy mode that allows for monoenergetic image reconstruction. For the GSI dual-energy scan, images were reconstructed at 140keV, the highest energy available; this level was chosen based on previous work that indicated that 140keV was more successful than lower energies at reducing metal artifacts.[68] The GSI images were reconstructed without and with the MARs algorithm ("GSI 140keV" and "MARs 140keV," respectively). In order to draw a fair comparison between the two different vendors, CT protocols were matched based on various acquisition and reconstruction parameters. Protocols for both the spine and dental phantoms are listed in Table 6.1. All reconstructed images were 12-bit depth images.

Table 6.1: Scan protocols for the spine and dental anthropomorphic phantoms using the Philips Brilliance and the GE Discovery CT750 HD scanners. All protocols are helical scans.

Phantom	Protocol	Pitch	mA	Tube rotation time (s)	Filter	Slice thickness (cm)	DFOV (cm)	Recon kernel	CTDIvol (mGy)
Spine phantom	Philips 120kVp	0.938	375	0.75	B	1.5	50	B	19.9
	GE 120kVp	0.969	440	0.5	Medium	1.25	50	Standard	20.3
	GE GSI preset #35	0.969	630	0.5	Medium	1.25	50	Standard	20.2
Dental phantom	Philips 120kVp	0.688	202	0.75	B	3.0	50	B	26.2
	GE 120kVp	0.516	375	0.6	Medium	2.5	50	Standard	26.0
	GE GSI preset #32	0.516	290	0.6	Medium	2.5	50	Standard	26.4

Treatment planning

For the two anthropomorphic phantoms, clinical treatment plans were created using the Pinnacle^3 v9.0 treatment planning system (Philips Healthcare, Andover, MA). For the spine case, a stereotactic spine radiosurgery (SSRS) treatment plan was created using nine posteriorly incident 6 MV step-and-shoot IMRT beams. The prescribed dose to the planning target volume was 6 Gy in a single fraction (with >90% coverage), and the resulting plan met the dose constraints for the spinal cord, heart, esophagus, skin, and whole lung according to IROC Houston's credentialing guidelines for the spine phantom. Figure 6.2 illustrates the location of the target, spinal cord, titanium rods, and the four ion chambers used to measure dose, as well as the isodose lines of the treatment plan for the spine case. For the dental case, a treatment plan was created using nine 6 MV step-and-shoot IMRT beams. Three target structures were created to represent high, intermediate, and low risk disease areas with prescription doses of 6.6, 6.0, and 5.4 Gy, respectively, in one fraction. The parotid glands, oral cavity, brainstem, and spinal cord were contoured and used as avoidance structures in the optimization process. Figure 6.3 illustrates the location of the three target structures, the parotid glands, the spinal cord, and the four ion chambers, as well as the isodose lines of the final treatment plan for the dental case. For both cases, plans were optimized using Pinnacle's inverse planning direct machine parameter optimization (DMPO) algorithm, and the dose variation across contoured ion chamber cavities was restricted to <2% standard deviation to ensure reliable measurements.

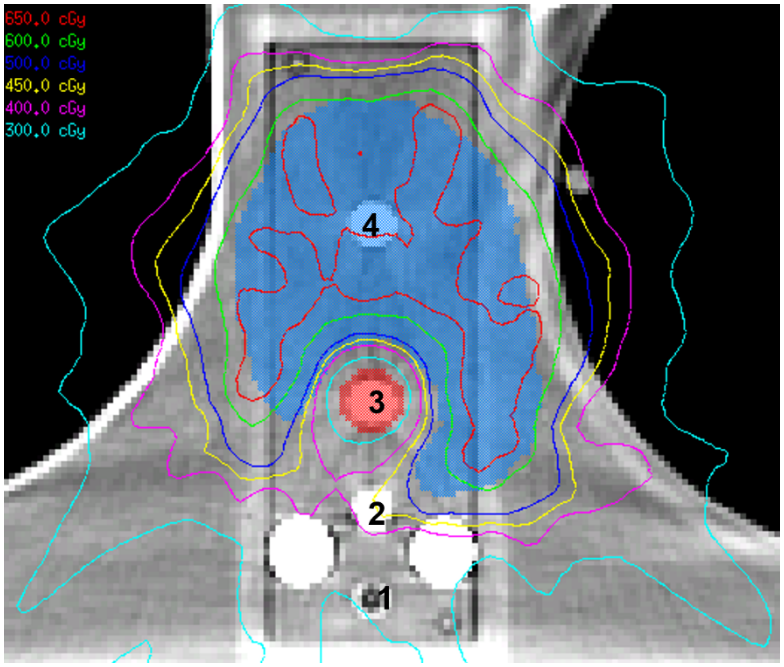


Figure 6.2: A CT image of the spine phantom illustrating the location of the two posterior titanium rods, the four ion chambers (numbered) used to measure dose at various distances from the titanium rods, the target structure (blue), and the spinal cord (red). The isodose lines for the resultant treatment plan are also shown.

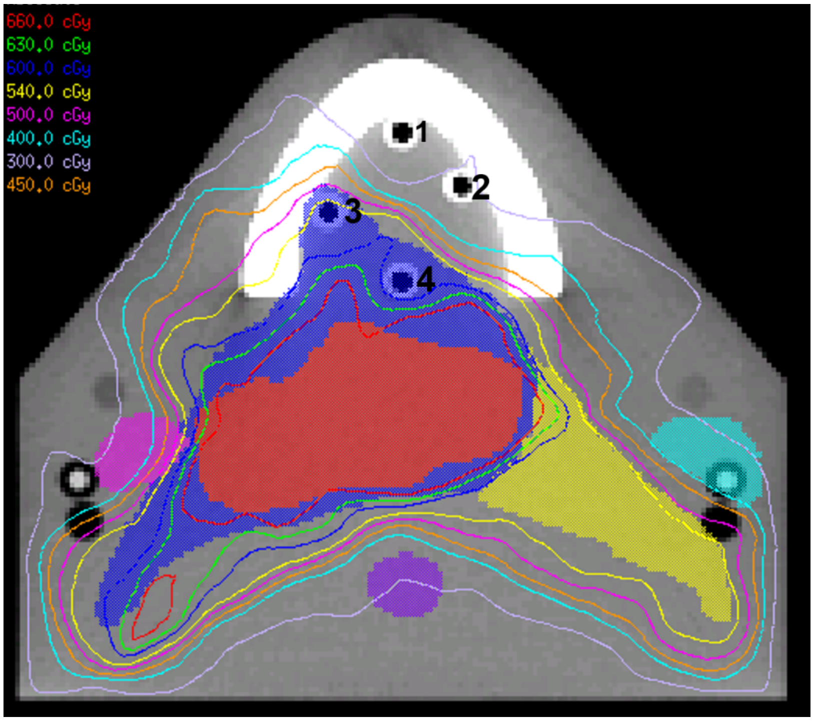


Figure 6.3: A CT image of the dental phantom (without fillings) illustrating the location of the three target structures (high, intermediate, and low risk indicated by red, blue, and yellow, respectively), the parotid glands (pink and light blue), spinal cord (purple), and the four ion chambers (numbered) used to measure dose at various locations in the oral cavity. The isodose lines for the resultant treatment plan are also shown.

Dose calculations

To investigate the impact of CT metal artifact reduction methods, the same treatment plan was copied to the various CT image sets acquired as described in section 2.2. All plans were then exported from Pinnacle to Mobius3D v1.3.1 (Mobius Medical Systems, Houston, TX). Mobius3D (M3D), a commercial treatment planning system verification tool that calculates dose using an independent collapsed cone C/S algorithm similar to Pinnacle's collapsed cone algorithm.[25] All dose calculations evaluated in this study were performed by the M3D algorithm.

The M3D system has reference or stock beam models for existing models of linear accelerators that can be customized using an automodeling feature. To better match our institution's Varian 2100 Clinac accelerator (Varian Associates, Palo, Alto, CA), beam model customization was performed by inputting measured dosimetry parameters (percent depth dose values for different depths and field sizes, off-axis factors, and output factors). This customized beam model was then validated with several verification plans, ranging in complexity from a simple water phantom with a single treatment beam to step-and-shoot IMRT clinical plans. The final, customized M3D beam model resulted in very good agreement with dose calculations from our institution's clinical Pinnacle treatment planning system, with all of our verification plans resulting in >96% of pixels passing for 3D gamma analysis at ±3%/3 mm (Appendix C).

Heterogeneity corrections for the M3D dose calculations were performed based on a CT calibration curve. This was complicated in this study because the polyenergetic CT images (120kVp) and the monochromatic dual-energy CT images (140keV) required two distinct CT number-to-density conversion curves with different slopes. The curves were created using a Gammex RMI 467 tissue characterization phantom, which contains a variety of tissue substitute materials of known density. However, because this study sought to assess the improvement in dose calculation accuracy due to different metal artifact reduction methods, we wanted the density mapping of the metal implants to be consistent across the different imaging techniques regardless of the different slopes in the two conversion curves.[61] Therefore, for all dose calculations, the saturated Hounsfield unit (HU) values in the images were mapped to the known density of the metals (4.51 g/cm^3 for titanium and 7.7 g/cm^3 for dental amalgam), essentially performing a density override for the metal region in the CT images. For our 12-bit CT images, the maximum HU value was 3071. However, the MARs algorithm consistently maps metals to a pre-defined HU level that is a function of the monochromatic reconstruction energy.[68] For 140keV GSI imaging with MARs, this value was approximately 1340 HU, so HU values greater than or equal to this threshold value in the MARs images were assigned to the true metal density. The majority of dose calculations performed in this study were conducted with this approach. However, in clinical practice, another approach to managing high-density materials is to cap the CT-to-density conversion curves at a maximum density of 1.82 g/cm^3 (corresponding to cortical bone). This approach prevents streak artifacts from being assigned to overly high density values. This capped-curve approach minimizes the dosimetric error caused by the artifacts but can also lead to an underestimation of attenuation caused by metal implants.[3] The dose calculation accuracy that can be achieved with this method was also evaluated for both phantoms using uncorrected 120kVp imaging.

To investigate the impact of implementing metal kernels for clinical dose calculations, two sets of metal kernels (titanium and silver) were implemented into M3D's collapsed cone C/S algorithm. Titanium kernels were used for the spine case, while silver kernels were used for the dental case. These metal kernels were generated using the EGSnrc Monte Carlo system and characterized in previous work.[65] The metal kernels were implemented in M3D by using a density threshold, i.e., for voxels with an assigned density greater than the threshold value (e.g., 4.0 g/cm^3), metal kernels were used to describe the energy deposition for energy released from those voxels. For voxels below the density threshold, water kernels were used. Kernel density scaling was still employed for voxels with density that did not exactly match one of the two kernel materials (water and metal).

M3D dose calculations were performed for both phantom cases (spine and dental) using a uniform 1.5 mm dose grid size. For each phantom, dose calculations were performed for uncorrected imaging ("baseline" calculation) and metal artifact reduction methods (O-MAR, GSI, and MARs). For each image set, dose calculations were performed with water kernels only and repeated with water and metal kernels. Thus, for each metal artifact reduction method studied, there were four dose calculations to compare in terms of accuracy (baseline calculation, baseline with metal kernels, artifact reduction calculation, and artifact reduction with metal kernels).

Phantom irradiation and analysis

Phantom irradiation was performed using a Varian Clinac 2100 linear accelerator. For each phantom, four Exradin A1SL ion chambers were used to measure the dose along with an axial EBT2 radiochromic film plane that intersected both the target structure and the metal implants. For each phantom, the treatment plan was delivered twice with the metal implant in place and twice with no metal implants to verify the reproducibility of the irradiation and measurements and to quantify the level of accuracy that can be achieved in the ideal case without metal.

For the ion chambers, raw measurement values were converted to dose using the ADCL-obtained calibration coefficient $N_{D,w}^{Co60}$, k_Q,[74] and correction factors P_{ion} and $P_{pol,}$ P_{elec}, and $P_{T,P.}$ Repeated ion chamber measurements were averaged. The error between calculated and measured dose was then reported for each chamber location. To put calculation errors in perspective, the level of accuracy achieved without metal implants was compared with the results obtained with the implants in place.

For the radiochromic film measurements, the optical density (OD) of the irradiated films was obtained using a CCD microdensitometer for radiochromic film (model CCD100; Photoelectron Corporation, Lexington, MA) approximately 4 days after irradiation. In order to convert the OD to absorbed dose, a calibration curve (third degree polynomial fit) was generated by irradiating films to absorbed doses ranging from 0.5 Gy to 15 Gy. To compare the measured planar dose against calculated dose, film registration and analysis were performed using a MATLAB-based software program developed at IROC Houston based on the known locations of pinpricks and landmarks in each phantom.[63] Once spatial registration was performed, the film-measured dose was normalized to perform relative 2D gamma analysis comparisons against calculated dose using a ±2%/2mm criteria and a region of interest chosen to highlight calculation differences near the metal implants. The ±2%/2mm criteria was chosen rather than the more clinically common ±3%/3mm criterion in order to better highlight dose differences due to artifact reduction methods and metal kernels.

6.3 Results

Spine phantom

The results for ion chamber measurements for the SSRS treatment plan delivered to the spine phantom with and without titanium rods are listed in Table 6.2. The reproducibility of repeat ion chamber measurements was within 0.5% for all measurement locations. The two chambers (1 and 2) that were closest to the titanium rods were most affected by the presence of the metal implants. Chambers 3 and 4, located in the spinal cord and target, respectively, were minimally affected by the presence of the rods, and dose calculation errors were <2.0% with or without metal. For the chambers affected by the metal, chamber 2 showed a decrease in measured dose (by 2.7%) caused by attenuation in the titanium rods, and both chambers 1 and 2 showed a decrease in calculation accuracy caused by the metal (increases in error by 0.7% and 0.5%, respectively). Therefore, we chose to focus our analysis on these two chamber locations.

Table 6.2: Ion chamber results for the spine and dental phantom (averaged over 2 repeat irradiations). For each chamber location, the ratio of the measured dose with metal implants in place to the dose measured with no metal is presented, along with the % error between calculated (Philips 120kVp uncorrected imaging) and measured doses.

Phantom	Ion Chamber	Ratio (with metal / no metal)	% Error		
			No metal	With metal	With metal*
Spine phantom	1	0.999	3.5%	4.2%	3.9%
	2	0.973	1.3%	1.8%	4.4%
	3	1.006	1.8%	1.6%	1.1%
	4	1.002	-0.8%	-0.7%	-0.7%
Dental phantom	1	0.999	3.0%	-11.5%	-1.1%
	2	1.004	1.7%	-7.5%	2.8%
	3	0.977	-0.9%	-10.2%	1.2%
	4	0.966	0.8%	-0.3%	7.2%

*Dose calculations were performed using a CT number-to-density curve that did not assign densities above that of cortical bone (1.82 g/cm^3).

For chambers 1 and 2, the percent error for dose calculations performed with metal artifact reduction and/or metal kernels is listed in Table 6.3. The artifact reduction methods and metal kernels generally did not have a large or even positive effect on dose calculation accuracy. In fact, O-MAR and GSI imaging typically caused small increases in error (although they were generally <1%). More drastically, the MARs algorithm notably decreased accuracy for both measurement locations. For chamber 2, MARs increased the error from 2.3% to 5.0% in comparison to uncorrected images. The MARs algorithm has previously been shown to overcorrect for the metal in CT images and cause an underestimation of the size of titanium implants;[43, 68] therefore, the observed overestimation in dose was likely the result of an underestimation of the size of the titanium rods in the MARs images and a subsequent underestimation of attenuation caused by the rods. Using metal kernels in the dose calculations also had a relatively small effect on accuracy (generally <0.5%), although improvements in accuracy were seen for chamber 2. In combining the two error-reduction strategies, some successes were found (particularly for chamber 2 with GSI imaging), but results were routinely comparable or slightly worse than using uncorrected 120kVp images with water kernels.

Table 6.3: Errors between calculated and ion chamber measured dose for the spine case. The baseline calculation error for uncorrected CT imaging is listed as well as the absolute change in calculation error when CT artifact reduction methods are used (O-MAR, GSI, and MARs), when metal kernels are used (MK), and when artifact reduction and metal kernels are used for the dose calculation.

	Chamber 1	Chamber 2	Average
Baseline % error*	4.3%	2.0%	3.2%
	Absolute change in % error		
O-MAR	+0.7%	+0.1%	+0.4%
GSI 140keV	+0.7%	-0.4%	+0.2%
MARs 140keV	+1.4%	+2.7%	+2.1%
MK	+0.1%	-0.3%	-0.1%
O-MAR+MK	+0.9%	-0.1%	+0.4%
GSI+MK	+0.8%	-0.5%	+0.2%
MARs+MK	+1.4%	+2.7%	+2.1%

*Average % error based on two uncorrected CT imaging methods (GE 120kVp and Philips 120kVp).

To spatially evaluate the agreement between calculated and measured dose, particularly in closer proximity to the titanium rods, gamma analysis was also performed for an axial film plane that intersected the titanium rods. Metal kernels affect the dose calculation near metal/tissue interfaces,[65] and the streak artifacts in the spine phantom were the most severe medial to the two titanium rods. Thus, we focused our analysis on the portion of the film in close proximity to the titanium rods (Figure 6.4a). For the two films irradiated with the titanium rods in the spine phantom, the average results are reported in Table 6.4. Although most of the changes in percentage of pixels passing due to the error-reduction strategies were not significant based on the reproducibility of our repeat films (approximately 5% pixels passing standard deviation), some of the more illustrative results are shown in Figure 6.4. Figure 6.4 has gamma index maps for irradiations performed with and without titanium rods. For the film measurements performed with titanium rods in place, GSI and MARs decreased passing rates, while metal kernels increased the passing rate. The largest change was seen for the MARs dose calculation, in which MARs resulted in a decrease in the percentage of pixels passing from 69.2% to 45.6%. These results broadly agree with our ion chamber results that showed small changes from the application of O-MAR or GSI, small changes (but improvement) from the application of metal kernels, and a substantial decrease in dose calculation accuracy from the application of MARs.

Table 6.4: The mean percentages of pixels passing for gamma analysis (±2%/2mm criteria) comparing film-measured dose to calculated dose for two repeated film measurements (with metal implants in place). Dose calculations were performed for the four possible combinations of the two error-reduction strategies (CT artifact reduction [AR] methods and metal kernels [MK]).

Phantom	AR method	Mean % pixels passing			
		Baseline	AR	Baseline +MK	AR+MK
Spine phantom	O-MAR	81.0	77.9	82.9	80.2
	GSI 140keV	74.3	62.4	80.7	60.3
	MARs 140keV	74.3	43.5	80.7	46.4
Dental phantom	O-MAR	69.8	72.0	70.5	71.8
	GSI 140keV	68.1	67.3	69.0	66.7
	MARs 140keV	68.1	71.6	69.0	70.4

*Baseline imaging was Philips 120kVp for O-MAR and GE 120kVp for both GSI 140keV and MARs 140keV imaging.

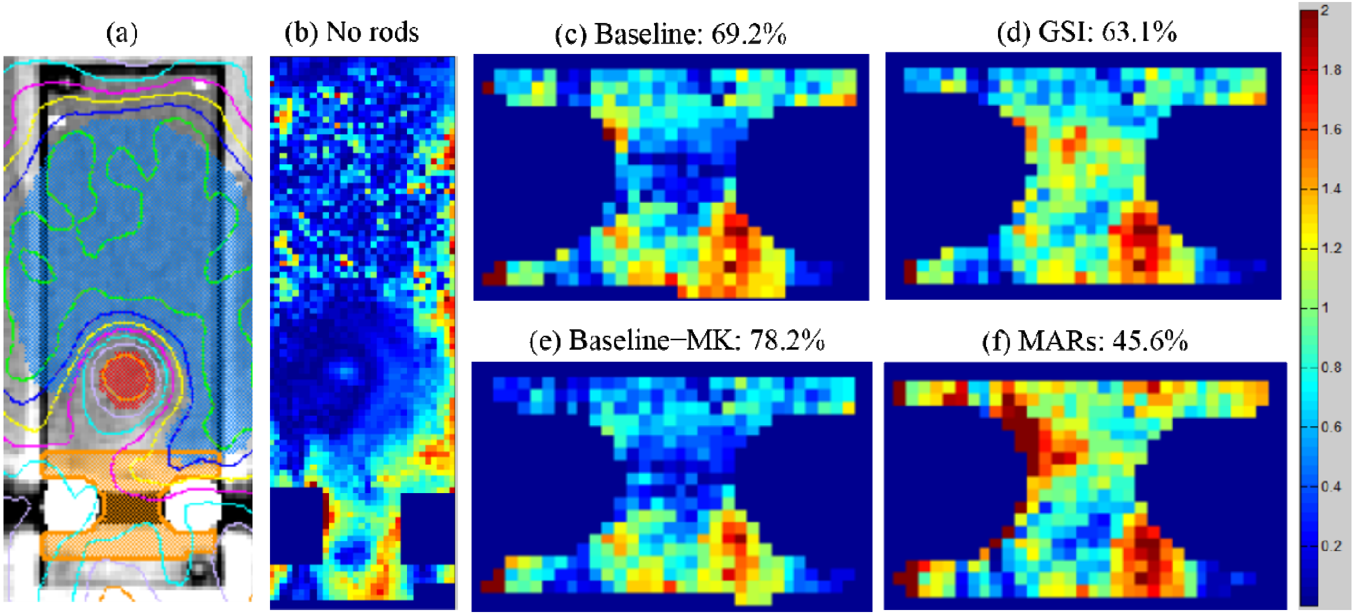


Figure 6.4: Film gamma analysis results for the spine case. Panel (a) shows the location of the region of interest (in orange) used for gamma analysis in panels (c)-(f), panel (b) shows the results for the entire film plane using ±3%/3mm criteria for an irradiation without metal rods (86.6% pixels passing), and panels (c)-(f) show the results for an irradiation with titanium rods using ±2%/2mm criteria for a sample of dose calculations using artifact reduction methods (GSI and MARs) and metal kernels (MK). The baseline dose calculation was performed using GE uncorrected 120kVp images.

Dental phantom

The ion chamber results for the base of tongue treatment plan delivered to the dental phantom with and without fillings are listed in Table 6.2. The reproducibility of repeat ion chamber measurements was within 0.7% for all measurement locations. In comparison to the spine phantom, the amalgam fillings had a greater effect on the measured dose, decreasing the dose by 2.3-3.4% in the target (chambers 3 and 4). Furthermore, the dose calculation accuracy was more affected by the presence of the fillings than the titanium spinal rods; the calculation errors were as high as 12%. Interestingly, for chamber 4, the measurement location in which the fillings decreased the dose the most, the calculation accuracy was very good (error < 1.0%) both with and without metal. In contrast, the chambers that were actually less affected by the fillings had much poorer calculation accuracy (errors >7.0%). Since all four chambers either had decreased measured dose or decreased calculation accuracy due to the metal fillings, all ion chamber locations were analyzed.

The percent error for various dose calculations performed with artifact reduction methods and/or metal kernels for the dental phantom with fillings is listed in Table 6.5. In comparison to baseline, uncorrected CT images, O-MAR was able to reduce errors for the three chamber locations that had the largest calculation errors. GSI imaging also generally reduced errors, though the gains in calculation accuracy were more modest than those observed with O-MAR. MARs was the most successful of the three methods for reducing calculation errors based on our ion chamber results, resulting in errors <4.0% for all four chamber locations. The success of the MARs algorithm can be explained by looking at the CT images of the dental phantom shown in Figure 6.5. In comparison to the other artifact reduction methods, MARs resulted in not only excellent artifact reduction but also better visualization of the size of the individual amalgam fillings (Figure 6.5e). Though O-MAR showed some artifact reduction (Figure 6.5c), residual artifacts still existed, making it difficult to discern the size of the fillings and the bone/metal and bone/tissue interfaces. Thus, using MARs images for the dose calculation, the attenuation caused by the fillings was more accurately calculated and not overestimated as much as with the other imaging methods. For the dental phantom, use of metal kernels in the dose calculation algorithm generally improved accuracy, though again the magnitude of the effect was generally small for our ion chamber measurements (≤0.6%). The largest effect due to metal kernels was seen for chamber 1, which was the most anteriorly located chamber.

Table 6.5: Errors between calculated and ion chamber measured dose for the dental case. The baseline calculation error for uncorrected CT imaging is listed as well as the absolute change in calculation error when CT artifact reduction methods are used (O-MAR, GSI, and MARs), when metal kernels are used (MK), and when artifact reduction and metal kernels are used for the dose calculation.

	Chamber 1	Chamber 2	Chamber 3	Chamber 4	Average
Baseline % error*	-10.8%	-6.2%	-9.2%	-0.8%	-6.7%
			Absolute change in % error		
O-MAR	-3.4%	-2.6%	-2.1%	+0.9%	-1.8%
GSI 140keV	-1.4%	-0.2%	+0.8%	-0.9%	-0.4%
MARs 140keV	-7.9%	-4.7%	-6.0%	+2.7%	-4.0%
MK	-0.6%	-0.2%	+0.1%	-0.0%	-0.2%
O-MAR+MK	-4.1%	-2.8%	-1.8%	+0.8%	-2.0%
GSI + MK	-2.0%	+0.1%	+0.9%	-0.9%	-0.4%
MARs + MK	-7.9%	-4.7%	-5.7%	+2.7%	-3.9%

*Average % error based on two uncorrected CT imaging methods (GE 120kVp and Philips 120kVp).

Table 6.4 shows the results of gamma analysis for a region of interest that contains the fillings, the bone material, and the oral cavity of the dental phantom (Figure 6.6a). This region of interest was chosen in order to highlight the area most affected by the streak artifacts and the region where metal kernels were most likely to affect the dose calculation. Though most of the changes in percentage of pixels passing were small, Figure 6.6 highlights areas in which artifact reduction improved the agreement between calculated and measured dose. The gamma index maps for both O-MAR and MARs showed improved agreement in comparison to uncorrected images, particularly anteriorly and near the teeth (Figures 6.6e and 6.6f). Furthermore, when the dose was calculated using uncorrected images, disagreement was seen at the locations of the streaks in the oral cavity (Figures 6.6c and 6.6d); these streaks of disagreement were improved when O-MAR and MARs were used for dose calculations.

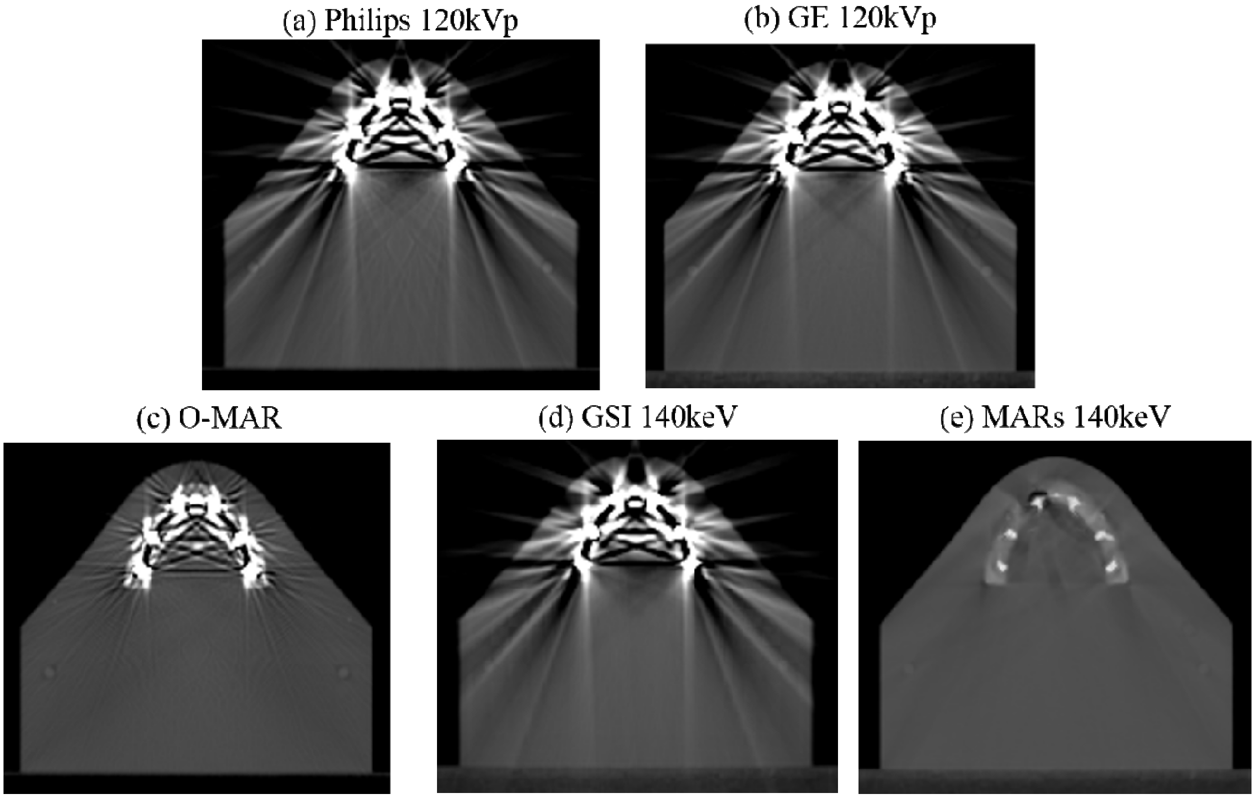


Figure 6.5: CT images of the dental phantom with amalgam fillings (window level = 500 and window width = 2500) using baseline, uncorrected imaging methods (Philips 120kVp and GE 120kVp) and artifact reduction methods (O-MAR, GSI 140keV, and MARs 140keV).

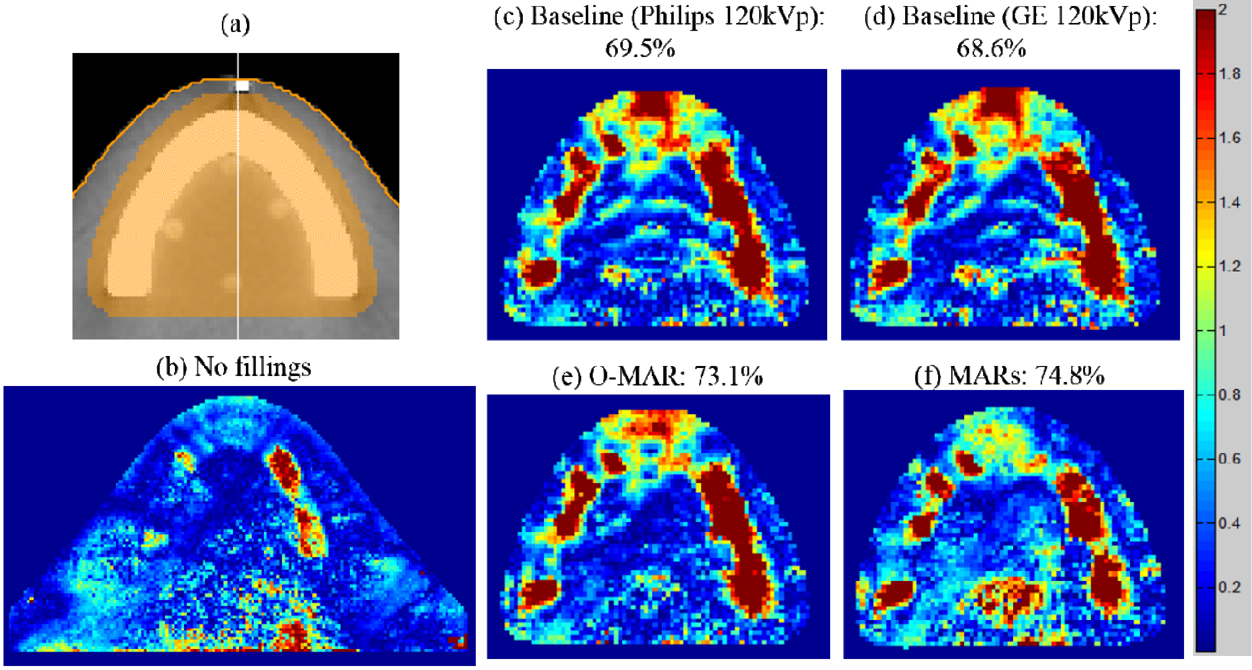


Figure 6.6: Film gamma analysis results for the dental phantom. Panel (a) shows the location of the region of interest (in orange) used for gamma analysis in panels (c)-(f), panel (b) shows results for the entire film plane using ±3%/3mm criteria for an irradiation without fillings (91.7% pixels passing), and

panels (c)-(f) show results using ±2%/2mm criteria for an irradiation with fillings for uncorrected, baseline images and artifact reduction methods (O-MAR and MARs).

Capped CT number-to-density conversion curve

For all the results presented up to this point, dose calculations were performed with a CT-to-density conversion method in which the saturated pixels in the image are mapped to the true density of the metal implant, essentially performing a density override for the metal. This was done as a possible clinical management solution in which the metal is known. It was also done so that density assignment for metal was consistent across the different imaging techniques in this study (polyenergetic 120kVp vs. monochromatic 140keV) to allow for a fair comparison. However, another clinical approach to minimizing the dosimetric impact of the streak artifacts was to use a CT number-to-density conversion curve that is capped at the density of cortical bone so that streaks do not result in overly inflated density assignments. Thus, this capped density approach can be considered an artifact mitigation method for dose calculation accuracy. Table 6.2 lists the calculation errors for our ion chamber results when this capped density approach was used for dose calculations for both phantoms. In comparison with the density override approach, the capped density curve resulted in poorer agreement for chamber 2 in the spine case, while the errors were similar for all other chambers. However, for the dental phantom, the capped density approach resulted in better agreement overall than the density override approach. The major exception to this improvement was for chamber 4, located in the target, in which the dose was overestimated by the capped density dose calculation (7.2% error), resulting because this method underestimates the attenuation caused by the metal fillings. Overall, when taking into account the capped density approach and the three CT artifact reduction methods, the MARs algorithm was still the most successful method for reducing calculation errors for the dental phantom.

6.4 Discussion

In this study, we investigated the use of CT metal artifact reduction methods and use of metal kernels in the dose calculation algorithm to improve dose calculation accuracy for two clinical scenarios: the case of a patient with spinal hardware receiving SSRS and the case of a patient with dental filling receiving head and neck radiation therapy. The high-Z dental amalgam fillings resulted in a larger dose perturbation, more severe imaging artifacts, and larger dose calculation errors than the low-Z titanium spinal rods. Unfortunately, no approach was found to be universally successful at managing both of the metal implants. Table 6.6 summarizes how CT metal artifact reduction methods and metal kernels affected dose calculation accuracy for each of the clinical cases.

Table 6.6: Summary of our results evaluating the two error-reduction strategies (CT metal artifact reduction methods and metal kernels) for clinical spine and dental cases. ✓ indicates that the strategy resulted in a small reduction in calculation errors, ✓✓ indicates a more substantial reduction in calculation errors, and ★ indicates that the strategy was the most successful strategy overall for reducing errors. X indicates that the strategy increased calculation errors, and – indicates that the strategy had little dosimetric impact.

	Spine	Dental	Comments
O-MAR	--	✓✓	Safest option for all-purpose CT simulation imaging
GSI 140keV	--	--	Not successful at reducing the severe artifacts caused by high Z dental fillings
MARs 140keV	X	★	Underestimation of metal size and distortion of metal shape can lead to increased calculation errors in comparison to uncorrected CT images
Metal Kernels	--	--	Reduces the calculation error by <1.0%

For the spine case, we found that the dose to the spinal cord was not affected by the titanium stabilization rods, in agreement with Liebross, et al.[70] nor was the dose to a measurement location in the target structure. However, the titanium rods did perturb the dose to two measurement locations closer to the rods (both approximately 8 mm from the edge of the titanium rods) and resulted in dose calculation errors of 1-4% for uncorrected CT images. This dose calculation inaccuracy could be particularly important for patients with titanium screws in close proximity to the target and spinal cord. O-MAR and GSI imaging had little effect on calculation accuracy, in agreement with a study by Spadea, et al.[12] that found that the dose differences between corrected and uncorrected images were not significant for a low-Z metal implant. In fact, for the SSRS treatment plan, the calculation errors were only increased slightly by the presence of the rods (in comparison to the case with no metal implants), indicating that limitations in small field beam modeling in the treatment planning system may dominate the error more than the imaging artifacts. Thus, the benefit of using a metal artifact reduction method such as O-MAR for the case of titanium spinal implants may be improved image quality and anatomical conspicuity rather than improved dose calculation accuracy. In contrast to GSI and O-MAR, the MARs algorithm resulted in a substantial increase in calculation errors, more than doubling the error in one measurement location (from 2.3% to 5.0%), due to an underestimation of the titanium rod diameter in the MARs images.[43, 68] Based on these data, it is not advisable to use the MARs algorithm with titanium implants because it can lead to an overestimation of dose downstream of metal and thus an overestimation of target coverage.

For the dental case, more severe streaks and larger calculation errors were observed (as high as 12%). This calculation error, which underestimates the dose to the oral cavity and the target, is likely due to a combination of imaging artifacts as well as the fillings appearing larger in the CT images than in reality. In contrast to the spine case, the MARs algorithm performed very well and reduced calculation errors by more than half at several locations in the oral cavity and the target. The success of MARs can be attributed to the fact that it corrected the severe artifacts caused by the dental amalgam and was able to more accurately represent the size and shape of the fillings than the other imaging methods. The O-MAR algorithm reduced streaks and improved accuracy for the dental case, though less so than MARs, while GSI monochromatic imaging had the smallest impact on accuracy of the three methods investigated and was the least successful at reducing the streaks caused by dental fillings. Film analysis revealed that MARs and O-MAR resulted in better agreement between measured and calculated doses, reducing errors associated with streak artifacts in the oral cavity. The case of a patient with dental fillings receiving radiation therapy to the head and neck is the clinical scenario for which metal artifact reduction methods can offer the most benefit, with target structures and avoidance structures clustered closely together, highly modulated and complex plans, and severe imaging artifacts. Reducing these imaging artifacts not only benefits the patient by improving anatomic conspicuity and reducing dose calculation errors in the affected region but can also reduce the hot and cold spots that can occur when treatment plans are created using uncorrected CT images; these hot and cold spots correspond to an increase in normal tissue complication probability and a decrease in tumor control probability.[75]

In summarizing the success of the artifact reduction methods in this study, the MARs algorithm resulted in the largest dosimetric impact of the three methods investigated, though the results were inconsistent. MARs increased calculation errors for the spine case while decreasing errors for the dental case. O-MAR was a more consistent artifact reduction method, either decreasing errors (dental case) or having little effect on calculation accuracy (spine case), although its benefit in the dental case was much less pronounced than for the MARs algorithm. O-MAR has been investigated in the literature for radiation therapy treatment planning without any findings indicating that it is detrimental for dose calculations.[37, 76] The last metal artifact reduction method, GSI imaging without MARs, had very little effect on calculation accuracy and was not successful in reducing artifacts associated with high-Z dental fillings. Though GSI dual-energy CT imaging may not be as successful as a dedicated artifact reduction algorithm, such as O-MAR or MARs, dual-energy CT nonetheless has many applications in radiation therapy aside from metal artifact reduction and still has the potential to improve the accuracy of treatments in other ways.[57, 66, 77] In summary, our results indicate that of the three methods investigated, O-MAR is the safest option for all-purpose metal artifact reduction in CT simulation imaging. MARs showed the greatest potential, but had serious limitations. MARs was developed primarily for diagnostic imaging and has not been fine-tuned for CT simulation imaging purposes. Thus, MARs should be used with great caution, especially for titanium implants, larger implants, and implants located near heterogeneities, as it can lead to metal distortion and introduce secondary artifacts.[68]

In addition to several commercial metal artifact reduction methods, we also investigated using a capped CT number-to-density conversion curve, i.e., a curve that maps CT numbers to a maximum density corresponding to cortical bone. In agreement with Wang, et al.[3] we found that for the spine case this approach resulted in a dose overestimation and increased error in comparison to performing a density override for the implant. For the dental fillings case, this capped curve approach prevented streak artifacts from being mapped to overly high densities and reduced calculation errors. However, it should be noted that this approach can overestimate target coverage for head and neck plans because it does not accurately take into account the attenuation due to the metal fillings.[78] Given little or no information about the composition of dental fillings, however, the capped curved approach may be reasonable.

In this study, we also investigated dose calculations performed with metal kernels. In comparison to dose calculations performed with water kernels only, metal kernels have been found to more accurately model the increased backscatter dose caused by metals.[65] Though we typically found some benefit from using metal kernels, the dosimetric impact was generally small. However, the maximum dosimetric benefit of metal kernels may not have been achieved in these clinical cases because the imaging artifacts resulted in messy and sometimes undiscernible metal/tissue interfaces and the many beam angles in these clinical intensity-modulated plans washed out the effects of the kernels to some extent. Despite the limited success of metal kernels in this study, there are clinical benefits to further pursuing more accurate calculation near metal implants. Modern treatment planning systems are not able to predict the backscatter dose enhancement due to metals.[3] For cases in which spinal hardware is close to the spinal cord, this limitation can lead to an underestimation of dose to the spinal cord, the dose-limiting structure for spinal treatments. For dental work, the local dose enhancement can lead to adverse reactions in the oral mucosa, and more accurate dose calculations could predict and prevent these painful oral complications.[6]

One limitation of the current study is that we investigated the use of metal artifact reduction and metal kernels for two specific clinical cases. Though these clinical cases are representative of patient cases commonly encountered in radiation therapy, all of our results may not be generalizable to other types of implants and treatment techniques. Nonetheless, our study provides valuable examples for when these error-reduction strategies are beneficial and when they may actually be detrimental.

6.5 Conclusions

Of the commercial metal artifact reduction methods investigated, O-MAR was the safest candidate for all-purpose artifact reduction for CT simulation imaging and treatment planning, as it either improved calculation accuracy (dental case) or had little dosimetric impact (spine case). Monochromatic imaging using dual-energy CT data (GSI imaging), though promising for other applications in radiation therapy, was not effective at reducing severe artifacts caused by dental fillings and had very little dosimetric impact for the two clinical cases in this study. The MARs algorithm yielded mixed results, dramatically improving accuracy for the dental case, but substantially worsening accuracy for the spinal case. The MARs algorithm can result in metal distortion, sometimes resulting in more accurate dimensions of metal implants and sometimes underestimating the size of implants. Due to its inconsistency, the MARs algorithm should be used with caution for CT simulation imaging, particularly for larger implants, implants near heterogeneities, and titanium implants. Implementing metal kernels in a commercial collapsed cone C/S algorithm resulted in small changes in calculation accuracy (<1.0%), generally improving accuracy. However, artifact reduction and accurate representation of the size of the implant in CT images was of greater dosimetric importance. For patients with metal implants, management of artifacts and accurate dose calculation remain challenging areas in need of improved solutions.

Chapter 7: Summary and Conclusions

7.1 General Summary and Conclusions

Two strategies for reducing dose calculation errors for patients with metal implants receiving external beam photon radiotherapy were investigated. The first of these two strategies, implementation of metal energy deposition kernels in the dose calculation algorithm, was investigated for its ability to better model metal/tissue interface effects and reduce errors associated with limitations of the dose calculation algorithm. Secondly, several commercial CT metal artifact reduction methods were investigated for their ability to reduce errors stemming for imaging artifacts caused by metal implants.

<u>*Metal Kernels*</u>

Metal energy deposition kernels, that describe the pattern of energy deposition when photons interact in metal rather than tissue, were generated and characterized. This study revealed that photon interactions in metal result in more lateral scatter, more backscatter, and less energy deposited in the forward direction than interactions in water. Thus, simply scaling the dimensions of water kernels based on density is not a good approximation for modeling photon interactions in metal, and this density scaling approximation get worse with increasing atomic number. We implemented metal kernels into a commercial collapsed cone convolution/superposition algorithm and evaluated the impact on accuracy for a simple phantom geometry with a rectangular metal implant and a single incident photon beam. In comparison to dose calculations with water kernels only, metal kernels resulted in better prediction of the backscatter enhancement at the proximal metal interface but decreased accuracy directly downstream of the metal. We also found that implementing metal kernels only had a local effect on the dose calculation, with the dosimetric impact of the metal kernels only extending about 1 cm from the metal interface.

Since metal kernels were not universally effective at improving accuracy, the dosimetric impact of using metal kernels for clinical treatment plans was small. With many beam angles in complex IMRT treatment plans, the dosimetric differences seen with the simple geometric phantom were somewhat washed out, and the impact of metal kernels was limited to about 0.5%. Furthermore, for clinical cases, CT imaging artifacts also decreased the efficacy of the metal kernels; the imaging artifacts resulted in indiscernible metal/tissue interfaces at times as well as interfaces sometimes located at the wrong location (due to the implant appearing larger than reality in the images). Therefore, the full potential of metal kernels for reducing dose calculation errors may not be fully realized without a very successful CT artifact reduction method. Though limited gains were observed for metal kernel dose calculations, the improved calculation of the backscatter dose enhancement upstream of metal implants was encouraging, and can potentially lead to better prediction of local dose enhancements due to dental restorations and prevention of adverse side effects in the oral cavity.

CT artifact reduction methods

We investigated several commercial artifact reduction methods for CT imaging, representing a range of technologies and approaches: the Philips O-MAR algorithm (a software-only projection modification algorithm), monochromatic imaging using dual-energy CT data (GSI), and combining a dedicated artifact correction algorithm with dual-energy CT monochromatic imaging (MARs). Each method was evaluated based on its effect on CT number accuracy, metal size accuracy, and ability to reduce the severity of artifacts using several common implants (hip prosthesis, spinal rods, and dental fillings). Furthermore, each method was also evaluated for how it affects dose calculation accuracy for two clinical treatment scenarios.

Of the three methods investigated, O-MAR was the safest candidate for all-purpose metal artifact reduction in CT simulation imaging. Though sometimes less successful than the MARs algorithm (e.g., hip prosthesis and dental fillings), O-MAR was much more consistent. For all types of implants investigated, it either reduced artifacts and improved calculation accuracy or had little effect on the calculation accuracy, but never worsened accuracy. It was successful in reducing artifacts caused by a large implant such as a hip prosthesis but also reduced errors for small implants such as dental fillings. Aside from introducing some secondary artifacts for the case of spinal rods in the thoracic region, we did not observe any behavior from O-MAR that would be undesirable for treatment planning (e.g., distortion of metal size).

In contrast, the MARs algorithm was not as consistent as O-MAR. We observed some of the most successful instances of artifact reduction by the MARs algorithm but also some serious failures. For instance, MARs was the most successful method for reducing artifacts due to a hip prosthesis, amalgam dental fillings, and a rectangular Cerrobend implant. However, the MARs algorithm was developed primarily for diagnostic imaging and thus has some behaviors that are not suited toward CT imaging for treatment planning. MARs images were shown to underestimate the size of titanium rods and to create low-density pockets inside large implants, such as a hip prosthesis. When metal distortion was observed in MARs images, it could substantially worsen dose calculation accuracy compared to using uncorrected CT images. Furthermore, for the heterogeneous head phantom used to evaluate the algorithm in Specific Aim 2, we found that MARs introduced artifacts into neighboring image slices that did not contain any metal. Based on these experiences with the MARs algorithm, it is clear that MARs is capable of reducing artifacts and improving calculation accuracy in certain scenarios, but should not be used as an all-purpose solution for CT simulation imaging, given the variety of implants encountered in radiation oncology. In fact, there are certain cases for which MARs may increase errors and thus extreme caution should be used:

- The patient has an implant composed of titanium. MARs underestimates the size of titanium implants, which can result in errors downstream of the implant.
- The patient has a large implant, such as a hip prosthesis, and you plan to treat with beams going through the prosthesis. MARs can create low-density pockets inside large implants, which also results in errors downstream of the implant.
- The implant is in a highly heterogeneous environment (dental fillings near a large air cavity). MARs can introduce secondary artifacts in this case, which may not have a large impact on calculation accuracy but can affect the ability to visualize the anatomy and contour the targets and surrounding normal structures.

Monochromatic GSI imaging, without any metal artifact reduction software applied, was also evaluated for its ability to reduce metal streak artifacts. Monochromatic GSI images have been shown to exhibit reduced beam hardening artifacts in comparison to polyenergetic imaging. However, our study indicates that GSI is only beneficial for smaller, low Z implants (such as titanium rods) and is not particularly successful at reducing artifacts for large or high Z implants (such as a hip prosthesis or dental fillings). For large, high Z implants that create more severe imaging artifacts, a dedicated artifact reduction method (such as O-MAR or MARs) is needed. Since it had the least success in reducing artifacts, GSI monochromatic imaging offered only a slight dosimetric benefit for clinical cases; GSI generally changed the accuracy of dose calculations by <1.0% at measurement locations in the target and organs at risk in comparison to uncorrected polyenergetic imaging (120kVp).

General comments on the management of radiation therapy patients with metal implants

Throughout this study, it was clear that achieving accurate dose calculations for patients with metal implants can be very challenging and that there are many sources of calculation errors aside from imaging artifacts and limitations of the dose calculation algorithm. For instance, we observed large dose calculation errors for both artifact-corrected and uncorrected CT images that resulted from an overestimation of the size of the implant in the images. For typical 12-bit CT images, the CT number saturates for most metals, and this CT number saturation can spread to adjacent pixels, resulting in a "blooming" effect that makes the metal implant appear larger in the images than in reality. Furthermore, saturated CT numbers gives no information about the composition or the electron density of the metal implant. In this study, we performed density overrides based on our knowledge of the implant materials. However, if a patient's medical record is incomplete, this information may not be available. Also, performing density overrides can be time-consuming and requires manual contouring of the implant on CT images that may overestimate the size of the implant. Given the complex nature of the problem, some general comments and recommendations are listed below:

- In order to accurately calculate the attenuation caused by a metal implant, some information about the density of the implant needs to be known. For large implants, such as a hip prosthesis, the attenuation can be substantial and not taking into account this attenuation can lead to large calculation errors and reduced target coverage; therefore, it is advisable to follow the recommendation of AAPM Task Group 63 to completely or partially avoid beam angles that enter through the prosthesis.[1]
- For smaller implants, such as dental fillings or spinal hardware, the attenuation due to the implant is smaller and it may not be possible to avoid beam arrangements that intersect the implant. In this case, if material composition information is available for the implant, a manual density override can be performed. If no material composition information is available or the implant can't be confidently delineated on the CT images, a capped CT number-to-density conversion curve approach is reasonable. A CT number-to-density conversion curve with a maximum density corresponding to cortical bone prevents

hyperintense streaks from being mapped to overly high densities and will minimize dose calculation errors associated with severe streak artifacts. However, it should be noted that this capped curve approach can lead to an underestimation of the attenuation caused by the metal implant and may overestimate target coverage.

- Use of metal artifact reduction algorithms is becoming increasingly common for CT simulation imaging in radiation oncology clinics. Users should be aware that these algorithms can sometimes result in unexpected behaviors (e.g., metal distortion and introduction of secondary artifacts in heterogeneous environments). Prior to adopting a commercial solution, an evaluation of the algorithm should be performed. This evaluation need not be as quantitative and thorough as the one performed in Specific Aim 2, but should involve a variety of implants, as the algorithms can behave differently depending on the type of metal, the size of the implant, and the surrounding environment.

7.2 Evaluation of the Hypothesis

The hypothesis for this project was that implementation of metal kernels in the convolution/superposition dose calculation algorithm, in conjunction with CT metal artifact reduction methods, can reduce calculation errors near metal implants by a factor of two. This hypothesis was evaluated in Specific Aim 3, which quantified the dosimetric impact of the two error-reduction strategies for both a simple geometric phantom and two clinical scenarios (using realistic implant geometries for spinal hardware and dental fillings and IMRT treatment plans). Although the hypothesis was not universally met, there were specific cases in which this factor of two error reduction was achieved. For a rectangular Cerrobend implant, use of the MARs algorithm reduced errors by greater than a factor of two both upstream and downstream of the implant, with or without use of metal kernels. For the clinical case of a patient with amalgam fillings receiving radiation therapy to the head and neck region, the MARs algorithm reduced errors by greater than a factor of two for three out of four measurement locations (the three locations in which the error was largest). For instance, for a measurement location located in one of the target structures, MARs reduced the local dose calculation error from 8.2% to 2.2%, nearly a factor of four reduction. Use of metal kernels typically had only a small impact on dose calculation accuracy, and a reduction of error by a factor of two could not be achieved with metal kernels alone. Though O-MAR was also successful at reducing calculation errors in some instances, it did not achieve the factor of two error reduction; the largest reduction in error by O-MAR was seen for the case of the dental fillings, where use of O-MAR and metal kernels together reduced the error at one of the measurement locations in the oral cavity by a factor of 1.6 (from 7.5% to 4.7%). In summary, though MARs was not universally successful at reducing calculation errors, and sometimes actually resulted in an increase in error, it did meet the hypothesis statement in certain scenarios (for Cerrobend and dental fillings) by reducing errors by at least a factor of two.

7.3 Future research and applications

CT metal artifact reduction methods are becoming increasingly popular in both diagnostic imaging and radiation oncology. Though this work only evaluated three commercial solutions, the methodology used to evaluate metal artifact reduction methods can be applied to other algorithms not included in this study. Furthermore, aside from external beam radiation therapy, the use of metal artifact reduction for CT imaging can be investigated for artifacts caused by brachytherapy applicators, which is an especially interesting application of artifact reduction since heterogeneous model-based dose calculations are becoming increasingly popular for brachytherapy treatments.[79] Also, another avenue of future work is investigating the use of these metal artifact reduction methods for proton therapy dose calculations. Proton therapy dose calculations are more sensitive to streak artifacts than photon dose calculations because the streak artifacts can result in proton range errors.[72, 73] Thus, the dosimetric benefits of using CT artifact reduction may be greater for protons than for the photon treatment plans evaluated in this study.

With dual-energy CT imaging, information about both the relative electron density and the effective atomic number of each voxel can be obtained. In comparison to polyenergetic CT imaging, this added information can lead to more accurate density assignment for photon dose calculations,[66] more accurate stopping power ratio assignment for proton dose calculations,[77] and improved tissue segmentation for Monte Carlo dose calculations.[57] In addition to improving dose calculation accuracy, dual-energy CT has many additional applications in radiation therapy, including virtual unenhanced CT simulation imaging and image guidance using dual-energy cone-beam CT.

Another source of dose calculation uncertainty for patients with metal implants is that limited information is available in typical CT images about the density and the material composition of the implant due to the saturation of CT numbers. Though not investigated in the current study, extended bit-depth CT images have the potential to be able to distinguish between stainless steel and titanium implants, allowing the user to make a more informed decision in performing manual density overrides.[73, 76, 80] As another avenue of future work, metal artifact reduction algorithms can be investigated for their success on extended bit-depth CT images.

Though use of metal energy deposition kernels was shown to only have a small effect on dose calculation accuracy in this study, the benefits of metal kernels may not have been fully realized. Our study highlighted that common commercial algorithms handle metal/tissue interfaces differently, and these subtle differences in the implementation of the collapsed cone algorithm can greatly affected calculation accuracy near the interface. Additional work on the details of implementing these metal kernels in collapsed cone convolution/superposition algorithms can be done to further improve accuracy. Furthermore, the benefits of metal kernels cannot be fully achieved without clean metal/tissue interfaces, so the success of metal kernels also hinges on the success of further improvements in CT metal artifact reduction.

Bibliography

1. C. Reft, R. Alecu, I. J. Das, B. J. Gerbi, P. Keall, E. Lief, B. J. Mijnheer, N. Papanikolaou, C. Sibata and J. Van Dyk, "Dosimetric considerations for patients with HIP prostheses undergoing pelvic irradiation. Report of the AAPM Radiation Therapy Committee Task Group 63," Med. Phys. **30**, 1162-1182 (2003).

2. C. H. Sibata, H. C. Mota, P. D. Higgins, D. Gaisser, J. P. Saxton and K. H. Shin, "Influence of hip prostheses on high energy photon dose distributions," Int J Radiat Oncol Biol Phys **18**, 455-461 (1990).

3. X. Wang, J. N. Yang, X. Li, R. Tailor, O. Vassilliev, P. Brown, L. Rhines and E. Chang, "Effect of spine hardware on small spinal stereotactic radiosurgery dosimetry," Phys Med Biol **58**, 6733-6747 (2013).

4. M. Pekmezci, B. Dirican, B. Yapici, M. Yazici, A. Alanay and S. Gurdalli, "Spinal implants and radiation therapy: the effect of various configurations of titanium implant systems in a single-level vertebral metastasis model," J. Bone Joint Surg. Am. **88**, 1093-1100 (2006).

5. M. H. Lin, J. Li, R. A. Price, Jr., L. Wang, C. C. Lee and C. M. Ma, "The dosimetric impact of dental implants on head-and-neck volumetric modulated arc therapy," Phys Med Biol **58**, 1027-1040 (2013).

6. D. W. Chin, N. Treister, B. Friedland, R. A. Cormack, R. B. Tishler, G. M. Makrigiorgos and L. E. Court, "Effect of dental restorations and prostheses on radiotherapy dose distribution: a Monte Carlo study," J. Appl. Clin. Med. Phys. **10**, 2853 (2009).

7. A. Boyer and E. Mok, "A photon dose distribution model employing convolution calculations," Med. Phys. **12**, 169-177 (1985).

8. T. R. Mackie, J. W. Scrimger and J. J. Battista, "A convolution method of calculating dose for 15-MV x rays," Med. Phys. **12**, 188-196 (1985).

9. A. Ahnesjo, P. Andreo and A. Brahme, "Calculation and application of point spread functions for treatment planning with high energy photon beams," Acta Oncol **26**, 49-56 (1987).

10. S. Spirydovich, L. Papiez, M. Langer, G. Sandison and V. Thai, "High density dental materials and radiotherapy planning: comparison of the dose predictions using superposition algorithm and fluence map Monte Carlo method with radiochromic film measurements," Radiother. Oncol. **81**, 309-314 (2006).

11. E. Wieslander and T. Knoos, "Dose perturbation in the presence of metallic implants: treatment planning system versus Monte Carlo simulations," Phys. Med. Biol. **48**, 3295-3305 (2003).

12. M. F. Spadea, J. M. Verburg, G. Baroni and J. Seco, "The impact of low-Z and high-Z metal implants in IMRT: A Monte Carlo study of dose inaccuracies in commercial dose algorithms," Med. Phys. **41**, 011702 (2014).

13. G. T. Herman, "Correction for beam hardening in computed tomography," Phys Med Biol **24**, 81-106 (1979).

14. J. F. Barrett and N. Keat, "Artifacts in CT: recognition and avoidance," Radiographics : a review publication of the Radiological Society of North America, Inc **24**, 1679-1691 (2004).

15. W. Kilby, J. Sage and V. Rabett, "Tolerance levels for quality assurance of electron density values generated from CT in radiotherapy treatment planning," Phys. Med. Biol. **47**, 1485-1492 (2002).

16. J. C. Chu, B. Ni, R. Kriz and V. Amod Saxena, "Applications of simulator computed tomography number for photon dose calculations during radiotherapy treatment planning," Radiother. Oncol. **55**, 65-73 (2000).

17. C. Glide-Hurst, D. Chen, H. Zhong and I. J. Chetty, "Changes realized from extended bit-depth and metal artifact reduction in CT," Medical Physics **40**, - (2013).

18. G. Wang, D. L. Snyder, J. A. O'Sullivan and M. W. Vannier, "Iterative deblurring for CT metal artifact reduction," IEEE transactions on medical imaging **15**, 657-664 (1996).

19. F. E. Boas and D. Fleischmann, "Evaluation of Two Iterative Techniques for Reducing Metal Artifacts in Computed Tomography," Radiology **259**, 894-902 (2011).

20. A. H. Mahnken, R. Raupach, J. E. Wildberger, B. Jung, N. Heussen, T. G. Flohr, R. W. Gunther and S. Schaller, "A new algorithm for metal artifact reduction in computed tomography: in vitro and in vivo evaluation after total hip replacement," Investigative radiology **38**, 769-775 (2003).

21. J. M. Verburg and J. Seco, "CT metal artifact reduction method correcting for beam hardening and missing projections," Physics in Medicine and Biology **57**, 2803 (2012).

22. E. Pessis, R. Campagna, J. M. Sverzut, F. Bach, M. Rodallec, H. Guerini, A. Feydy and J. L. Drape, "Virtual monochromatic spectral imaging with fast kilovoltage switching: reduction of metal artifacts at CT," Radiographics : a review publication of the Radiological Society of North America, Inc **33**, 573-583 (2013).

23. B. Fraass, K. Doppke, M. Hunt, G. Kutcher, G. Starkschall, R. Stern and J. Van Dyke, "American Association of Physicists in Medicine Radiation Therapy Committee Task Group 53: quality assurance for clinical radiotherapy treatment planning," Med. Phys. **25**, 1773-1829 (1998).

24. P. E. Metcalfe, P. W. Hoban, D. C. Murray and W. H. Round, "Beam hardening of 10 MV radiotherapy x-rays: analysis using a convolution/superposition method," Phys. Med. Biol. **35**, 1533-1549 (1990).

25. N. Childress, E. Stephens, D. Eklund and M. Zhang, "Mobius3D White Paper: Dose Calculation Algorithm Rev. 0," Mobius Medical Systems, LP (2012).

26. T. McNutt, "Pinnacle3 White Paper on Dose Calculations," (2002).

27. A. Ahnesjo, "Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media," Med. Phys. **16**, 577-592 (1989).

28. J. M. Lydon, "Photon dose calculations in homogeneous media for a treatment planning system using a collapsed cone superposition convolution algorithm," Phys. Med. Biol. **43**, 1813-1822 (1998).

29. T. R. Mackie, A. F. Bielajew, D. W. Rogers and J. J. Battista, "Generation of photon energy deposition kernels using the EGS Monte Carlo code," Phys. Med. Biol. **33**, 1-20 (1988).

30. H. H. Liu, T. R. Mackie and E. C. McCullough, "Correcting kernel tilting and hardening in convolution/superposition dose calculations for clinical divergent and polychromatic photon beams," Med. Phys. **24**, 1729-1741 (1997).

31. P. W. Hoban, D. C. Murray and W. H. Round, "Photon beam convolution using polyenergetic energy deposition kernels," Phys. Med. Biol. **39**, 669-685 (1994).

32. N. Papanikolaou, T. R. Mackie, C. Meger-Wells, M. Gehring and P. Reckwerdt, "Investigation of the convolution method for polyenergetic spectra," Med. Phys. **20**, 1327-1336 (1993).

33. J. E. O'Connor, "The density scaling theorem applied to lateral electronic equilibrium," Med. Phys. **11**, 678-680 (1984).

34. M. K. Woo and J. R. Cunningham, "The validity of the density scaling method in primary electron transport for photon and electron beams," Med. Phys. **17**, 187-194 (1990).

35. J. T. Bushberg, *The essential physics of medical imaging*, 3rd ed. (Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, 2012).

36. "Metal Artifact Reduction for Orthopedic Implants (O-MAR) White Paper," Philips Healthcare (Philips CT Clinical Science 2012).

37. H. Li, C. Noel, H. Chen, H. Harold Li, D. Low, K. Moore, P. Klahr, J. Michalski, H. A. Gay, W. Thorstad and S. Mutic, "Clinical evaluation of a commercial orthopedic metal artifact reduction tool for CT simulations in radiation therapy," Med. Phys. **39**, 7507-7517 (2012).

38. L. Yu, S. Leng and C. H. McCollough, "Dual-energy CT-based monochromatic imaging," Am. J. Roentgenol. **199**, S9-S15 (2012).

39. R. K. Kaza, J. F. Platt and A. J. Megibow, "Dual-energy CT of the urinary tract," Abdominal imaging **38**, 167-179 (2013).

40. A. C. Silva, B. G. Morse, A. K. Hara, R. G. Paden, N. Hongo and W. Pavlicek, "Dual-energy (spectral) CT: applications in abdominal imaging," Radiographics **31**, 1031-1046; discussion 1047-1050 (2011).

41. W. A. Kalender, W. H. Perman, J. R. Vetter and E. Klotz, "Evaluation of a prototype dual-energy computed tomographic apparatus. I. Phantom studies," Medical Physics **13**, 334-339 (1986).

42. M. M. Goodsitt, E. G. Christodoulou and S. C. Larson, "Accuracies of the synthesized monochromatic CT numbers and effective atomic numbers obtained with a rapid kVp switching dual energy CT scanner," Med. Phys. **38**, 2222-2232 (2011).

43. Y. H. Lee, K. K. Park, H. T. Song, S. Kim and J. S. Suh, "Metal artefact reduction in gemstone spectral imaging dual-energy CT with and without metal artefact reduction software," European radiology **22**, 1331-1340 (2012).

44. Y. Wang, B. Qian, B. Li, G. Qin, Z. Zhou, Y. Qiu, X. Sun and B. Zhu, "Metal artifacts reduction using monochromatic images from spectral CT: evaluation of pedicle screws in patients with scoliosis," European journal of radiology **82**, e360-366 (2013).

45. O. R. Brook, S. Gourtsoyianni, A. Brook, A. Mahadevan, C. Wilcox and V. Raptopoulos, "Spectral CT with metal artifacts reduction software for improvement of tumor visibility in the vicinity of gold fiducial markers," Radiology **263**, 696-705 (2012).

46. D. W. O. Rogers, I. Kawrakow, J. P. Seuntjens, B. R. B. Walters, and E. Mainegra-Hing, "NRC User Codes for EGSnrc," Technical Report PIRS-702(RevB), National Research Council of Canada, Ottowa, Canada (2003).

47. D. W. O. Rogers, B. Walters, and I. Kawrakow, "BEAMnrc Users Manual," NRCC Report PIRS-0509(A)revK, National Research Council of Canada, Ottawa, Canada (2009).

48. S. H. Cho, O. N. Vassiliev, S. Lee, H. H. Liu, G. S. Ibbott and R. Mohan, "Reference photon dosimetry data and reference phase space data for the 6 MV photon beam from varian clinac 2100 series linear accelerators," Med. Phys. **32**, 137-148 (2005).

49. O. N. Vassiliev, U. Titt, S. F. Kry, F. Ponisch, M. T. Gillin and R. Mohan, "Monte Carlo study of photon fields from a flattening filter-free clinical accelerator," Med. Phys. **33**, 820-827 (2006).

50. ICRU, "Stopping powers for electrons and positrons," ICRU Report 37, ICRU, Washington D.C. (1984).

51. W. C. Roesch, "Dose for non-electronic equilibrium conditions," Radiat Res **9**, 399-410 (1958).

52. E. Mainegra-Hing, D. W. Rogers and I. Kawrakow, "Calculation of photon energy deposition kernels and electron dose point kernels in water," Med. Phys. **32**, 685-699 (2005).

53. I. Kawrakow, Rogers, E. Mainegra-Hing, D. W. O, F. Tessier, and B. R. B. Walters, "The EGSnrc Code System: Monte Carlo Simulation of Electron and Photon Transport," Report PIRS-701, National Research Council of Canada, Ottowa, Canada (2011).

54. N. Papanikolaou, J. J. Battista and A. Boyer, “Tissue inhomogeneity corrections for megavoltage photon beams AAPM Report No. 85”, Medical Physics Publishing, Madison WI (2004).

55. S. Zhao, D. D. Robertson, G. Wang, B. Whiting and K. T. Bae, "X-ray CT metal artifact reduction using wavelets: an application for imaging total hip prostheses," IEEE transactions on medical imaging **19**, 1238-1247 (2000).

56. J. Wei, G. A. Sandison, W. C. Hsi, M. Ringor and X. Lu, "Dosimetric impact of a CT metal artefact suppression algorithm for proton, electron and photon therapies," Phys. Med. Biol. **51**, 5183-5197 (2006).

57. M. Bazalova, L. Beaulieu, S. Palefsky and F. Verhaegena, "Correction of CT artifacts and its influence on Monte Carlo dose calculations," Med. Phys. **34**, 2119-2132 (2007).

58. G. Hilgers, T. Nuver and A. Minken, "The CT number accuracy of a novel commercial metal artifact reduction algorithm for large orthopedic implants," J. Appl. Clin. Med. Phys. **15**, 4597 (2014).

59. J. Hsieh, "TU-E-210A-01: Dual-Energy CT with Fast-KVp Switch," Medical Physics **36**, 2749-2749 (2009).

60. B. Li, G. Yadava and J. Hsieh, "Quantification of head and body CTDI(VOL) of dual-energy x-ray CT with fast-kVp switching," Med Phys **38**, 2595-2601 (2011).

61. M. Yagi, T. Ueguchi, M. Koizumi, T. Ogata, S. Yamada, Y. Takahashi, I. Sumida, Y. Akino, K. Konishi, F. Isohashi, N. Tomiyama, Y. Yoshioka and K. Ogawa, "Gemstone spectral imaging: determination of CT to ED conversion curves for radiotherapy treatment planning," J. Appl. Clin. Med. Phys. **14**, 4335 (2013).

62. D. Zhang, X. Li and B. Liu, "Objective characterization of GE discovery CT750 HD scanner: gemstone spectral imaging mode," Med Phys **38**, 1178-1188 (2011).

63. D. S. Followill, D. R. Evans, C. Cherry, A. Molineu, G. Fisher, W. F. Hanson and G. S. Ibbott, "Design, development, and implementation of the radiological physics center's pelvis and thorax anthropomorphic quality assurance phantoms," Med. Phys. **34**, 2070-2076 (2007).

64. D. Prell, Y. Kyriakou, M. Beister and W. A. Kalender, "A novel forward projection-based metal artifact reduction method for flat-detector computed tomography," Physics in Medicine and Biology **54**, 6575 (2009).

65. J. Y. Huang, D. Eklund, N. L. Childress, R. M. Howell, D. Mirkovic, D. S. Followill and S. F. Kry, "Investigation of various energy deposition kernel refinements for the convolutionsuperposition method," Med. Phys. **40**, 121721 (2013).

66. M. Tsukihara, Y. Noto, R. Sasamoto, T. Hayakawa and M. Saito, "Initial implementation of the conversion from the energy-subtracted CT number to electron density in tissue inhomogeneity

corrections: An anthropomorphic phantom study of radiotherapy treatment planning," Med. Phys. **42**, 1378 (2015).

67. C. Glide-Hurst, D. Chen, H. Zhong and I. J. Chetty, "Changes realized from extended bit-depth and metal artifact reduction in CT," Med. Phys. **40**, 061711 (2013).

68. J. Y. Huang, J. R. Kerns, J. L. Nute, X. Liu, P. A. Balter, F. C. Stingo, D. S. Followill, D. Mirkovic, R. M. Howell and S. F. Kry, "An evaluation of three commercially available metal artifact reduction methods for CT imaging," Phys. Med. Biol. **60**, 1047-1067 (2015).

69. S. H. Son, Y. N. Kang and M. R. Ryu, "The effect of metallic implants on radiation therapy in spinal tumor patients with metallic spinal implants," Med. Dosim. **37**, 98-107 (2012).

70. R. H. Liebross, G. Starkschall, P. F. Wong, J. Horton, Z. L. Gokaslan and R. Komaki, "The effect of titanium stabilization rods on spinal cord radiation dose," Med. Dosim. **27**, 21-24 (2002).

71. N. Suchowerska, P. Hoban, M. Butson, A. Davison and P. Metcalfe, "Directional dependence in film dosimetry: radiographic and radiochromic film," Phys. Med. Biol. **46**, 1391-1397 (2001).

72. W. D. Newhauser, A. Giebeler, K. M. Langen, D. Mirkovic and R. Mohan, "Can megavoltage computed tomography reduce proton range uncertainties in treatment plans for patients with large metal implants?," Phys. Med. Biol. **53**, 2327-2344 (2008).

73. M. Axente, A. Paidi, R. Von Eyben, C. Zeng, A. Bani-Hashemi, A. Krauss and D. Hristov, "Clinical evaluation of the iterative metal artifact reduction algorithm for CT simulation in radiotherapy," Med. Phys. **42**, 1170 (2015).

74. M. McEwen, L. DeWerd, G. Ibbott, D. Followill, D. W. Rogers, S. Seltzer and J. Seuntjens, "Addendum to the AAPM's TG-51 protocol for clinical reference dosimetry of high-energy photon beams," Med. Phys. **41**, 041501 (2014).

75. Y. Kim and W. A. Tome, "On the radiobiological impact of metal artifacts in head-and-neck IMRT in terms of tumor control probability (TCP) and normal tissue complication probability (NTCP)," Medical & biological engineering & computing **45**, 1045-1051 (2007).

76. C. Glide-Hurst, D. Chen, H. Zhong and I. J. Chetty, "Changes realized from extended bit-depth and metal artifact reduction in CT," Med. Phys. **40**, 061711 (2013).

77. N. Hunemohr, B. Krauss, C. Tremmel, B. Ackermann, O. Jakel and S. Greilich, "Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates," Phys. Med. Biol. **59**, 83-96 (2014).

78. N. Mail, Y. Albarakati, M. Ahmad Khan, F. Saeedi, N. Safadi, S. Al-Ghamdi and A. Saoudi, "The impacts of dental filling materials on RapidArc treatment planning and dose delivery: challenges and solution," Med. Phys. **40**, 081714 (2013).

79. L. Beaulieu, A. Carlsson Tedgren, J. F. Carrier, S. D. Davis, F. Mourtada, M. J. Rivard, R. M. Thomson, F. Verhaegen, T. A. Wareing and J. F. Williamson, "Report of the Task Group 186 on model-based dose calculation methods in brachytherapy beyond the TG-43 formalism: current status and recommendations for clinical implementation," Med. Phys. **39**, 6208-6236 (2012).

80. C. Coolens and P. J. Childs, "Calibration of CT Hounsfield units for radiotherapy treatment planning of patients with metallic hip prostheses: the use of the extended CT-scale," Phys. Med. Biol. **48**, 1591-1603 (2003).

Appendix A: Supplementary kernel data

In chapter 3, various energy deposition kernel refinements for the convolution/superposition dose calculation method were investigated. One such refinement was the calculation of high-resolution energy deposition kernels (i.e., kernels generated with a greater number of angular and radial sample points). High-resolution water kernels were generated as described in Section 3.2 for several monoenergetic photon energies. These high-resolution water kernels were then characterized by calculating the total energy fraction (F_{tot}) and primary energy fraction (F_{prim}), as well as the effective depth of penetration $\bar{z}$, effective radial distance $\bar{r}$, and effective lateral distance $\bar{y}$ (equation 3.1) for each kernel energy, and the result are listed in Table A.1.

In addition to comparing water kernel dose calculations against calculations performed with material-specific kernels (e.g., water and titanium kernels), we also performed EGSnrc Monte Carlo simulations for simple metal implant geometries. To perform calculations for our institution's Varian 2100 Clinac linear accelerator, the Monte Carlo package BEAMnrc was used to create an accelerator model consisting of the target, primary collimator, flattening filter, and jaws, and this model was used to generate particle phase space data of radiation exiting the accelerator. With this phase space data, the DOSXYZnrc package could then be used to calculate dose for a 10 cm x 10 cm field size 6-MV photon beam incident on a simple titanium implant geometry (30 x 30 x 30 cm^3 water phantom with a 4 x 4 x 4 cm^3 titanium insert embedded in it). Dose calculations performed with Mobius3D (M3D) with water kernels only and with water and titanium kernels were then compared to the Monte Carlo results (Figure A.1). With titanium kernels, better accuracy and better agreement with Monte Carlo results are achieved in the upstream region. In comparison to water kernels only, metal kernels result in better modeling of the backscatter dose enhancement. However, metal kernels result in decreased accuracy directly downstream of the metal implant. In this region, the M3D dose calculation with water kernels only agrees most closely to Monte Carlo. These results further support the film results obtained in Chapter 5.

Table A.1: The total energy fraction (F_{tot}), primary energy fraction (F_{prim}), effective depth of penetration $\bar{z}$, effective radial distance $\bar{r}$, and effective lateral distance $\bar{y}$ for high resolution, water-based energy deposition kernels for monoenergetic incident photons of various energies. In parenthesis are the percent uncertainties (one standard deviation).

Energy (MeV)	F_{tot}	F_{prim}	$\bar{z}$ (cm)	$\bar{r}$ (cm)	$\bar{y}$ (cm)
0.10	0.9981 (0.0053)	0.1483 (0.0052)	7.800E-03	1.250E-02	9.800E-03
0.15	0.9953 (0.0043)	0.1831 (0.0041)	8.000E-03	1.250E-02	9.700E-03
0.20	0.9929 (0.0039)	0.2161 (0.0038)	8.100E-03	1.260E-02	9.700E-03
0.30	0.9890 (0.0037)	0.2689 (0.0036)	8.500E-03	1.300E-02	9.800E-03
0.40	0.9859 (0.0036)	0.3088 (0.0034)	1.420E-02	2.060E-02	1.490E-02
0.50	0.9831 (0.0036)	0.3404 (0.0032)	2.190E-02	3.110E-02	2.200E-02
0.60	0.9807 (0.0036)	0.3666 (0.0030)	3.140E-02	4.380E-02	3.060E-02
0.80	0.9763 (0.0036)	0.4076 (0.0028)	5.380E-02	7.360E-02	5.020E-02
1.00	0.9725 (0.0035)	0.4390 (0.0027)	7.990E-02	0.1073	7.160E-02
1.25	0.9682 (0.0035)	0.4694 (0.0025)	0.1155	0.1522	9.900E-02
1.50	0.9642 (0.0034)	0.4928 (0.0024)	0.1527	0.1989	0.1263
2.00	0.9574 (0.0033)	0.5280 (0.0023)	0.2958	0.2958	0.1798
3.00	0.9473 (0.0031)	0.5751 (0.0022)	0.4077	0.4936	0.2782
4.00	0.9406 (0.0029)	0.6075 (0.0021)	0.5842	0.6887	0.3646
5.00	0.9362 (0.0027)	0.6323 (0.0021)	0.7628	0.8817	0.4422
6.00	0.9334 (0.0026)	0.6523 (0.0020)	0.9403	1.071	0.5121
8.00	0.9306 (0.0024)	0.6834 (0.0020)	1.286	1.433	0.6318
10.0	0.9298 (0.0023)	0.7063 (0.0019)	1.623	1.781	0.7336
15.0	0.9308 (0.0020)	0.7430 (0.0018)	2.427	2.602	0.9386
20.0	0.9322 (0.0019)	0.7630 (0.0018)	3.178	3.362	1.097
30.0	0.9312 (0.0018)	0.7769 (0.0017)	4.584	4.776	1.340
40.0	0.9260 (0.0017)	0.7745 (0.0017)	5.861	6.054	1.518

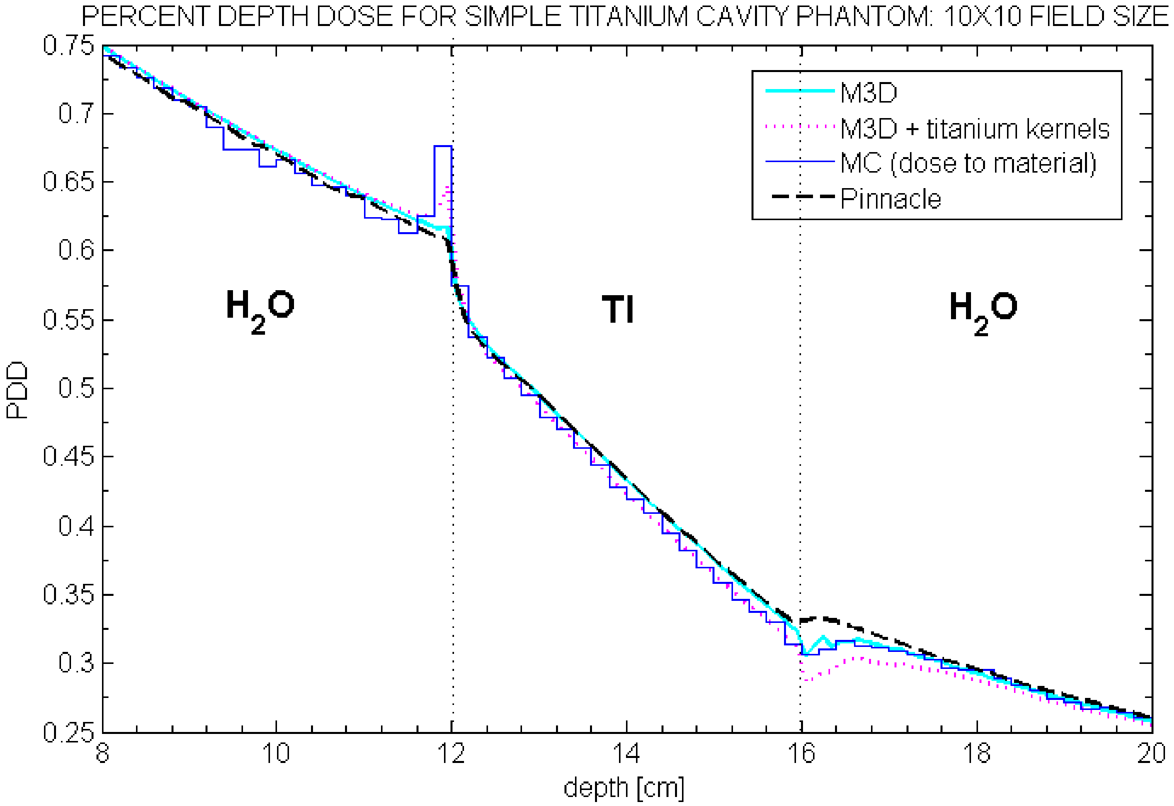


Figure A.1: Percent depth dose curves calculated using water kernels only ("M3D"), with water and titanium kernels, EGSnrc Monte Carlo simulation ("MC"), and Pinnacle's collapsed cone convolution algorithm.

Appendix B: Supplementary metal artifact reduction data

In chapter 4, several commercial metal artifact reduction methods were evaluated based on criteria relevant to treatment planning and dose calculation accuracy. To evaluate CT number accuracy, the RMI 467 tissue characterization phantom (Gammex, Middleton, WI) was scanned using different imaging techniques (uncorrected 120kVp imaging and metal artifact reduction methods), and the resulting images were analyzed. To quantify CT number accuracy, HU errors ($\overline{\Delta HU}$ as defined by equation 4.1) were calculated for select tissue substitute regions of interest. For a phantom configuration with a unilateral aluminum and a unilateral Cerrobend insert, the results are shown in Figure B.1. The mean CT number errors caused by the aluminum insert/plug were small ($\overline{\Delta HU} < 20$ HU), while the errors caused by the Cerrobend plug were larger and very similar to those seen with stainless steel.

While chapter 4 focused on investigating two specific monochromatic energies for GSI imaging (70keV and 140keV), we also investigated how the CT number error changes as a function of reconstruction energy. For all configurations of the RMI phantom (with different unilateral metal inserts and also a bilateral configuration), $\overline{\Delta HU}$ was calculated for select tissue substitute regions of interest, and the results are shown in Figures B.2 and B.3. In general, we found that the HU error decreases with increasing reconstruction energy for GSI images with and without the MARs algorithm applied. In investigating GSI imaging for dose calculations in chapters 5 and 6, we thus chose to use the highest reconstruction energy available (140keV).

Additionally, for the metal implants included in chapter 4 (Co-Cr hip prosthesis, stainless steel, and titanium), the CT number of the metal itself was also investigated for GSI imaging. Unlike conventional polyenergetic images that have saturated CT values for most metals, the GSI images showed some non-saturated HU values at lower reconstruction energies (Figure B.4a). When the MARs algorithm was applied, the metal HU values are consistently mapped to

HU values that vary as a function of reconstruction energy but are independent of the type of metal (Figure B.4b). This data suggests that metal pixels are identified prior to application of the MARs algorithm, and then these metal pixels are replaced by pre-defined values after the algorithm has been applied to the image.

In our evaluation of the MARs algorithm, we found that application of MARs could result in underestimation of the size of titanium rods and cause metal distortion. This behavior was observed for scans of a hip prosthesis, in which the MARs algorithm caused the most superior aspect of the prosthesis to no longer be visible (Figure B.5).

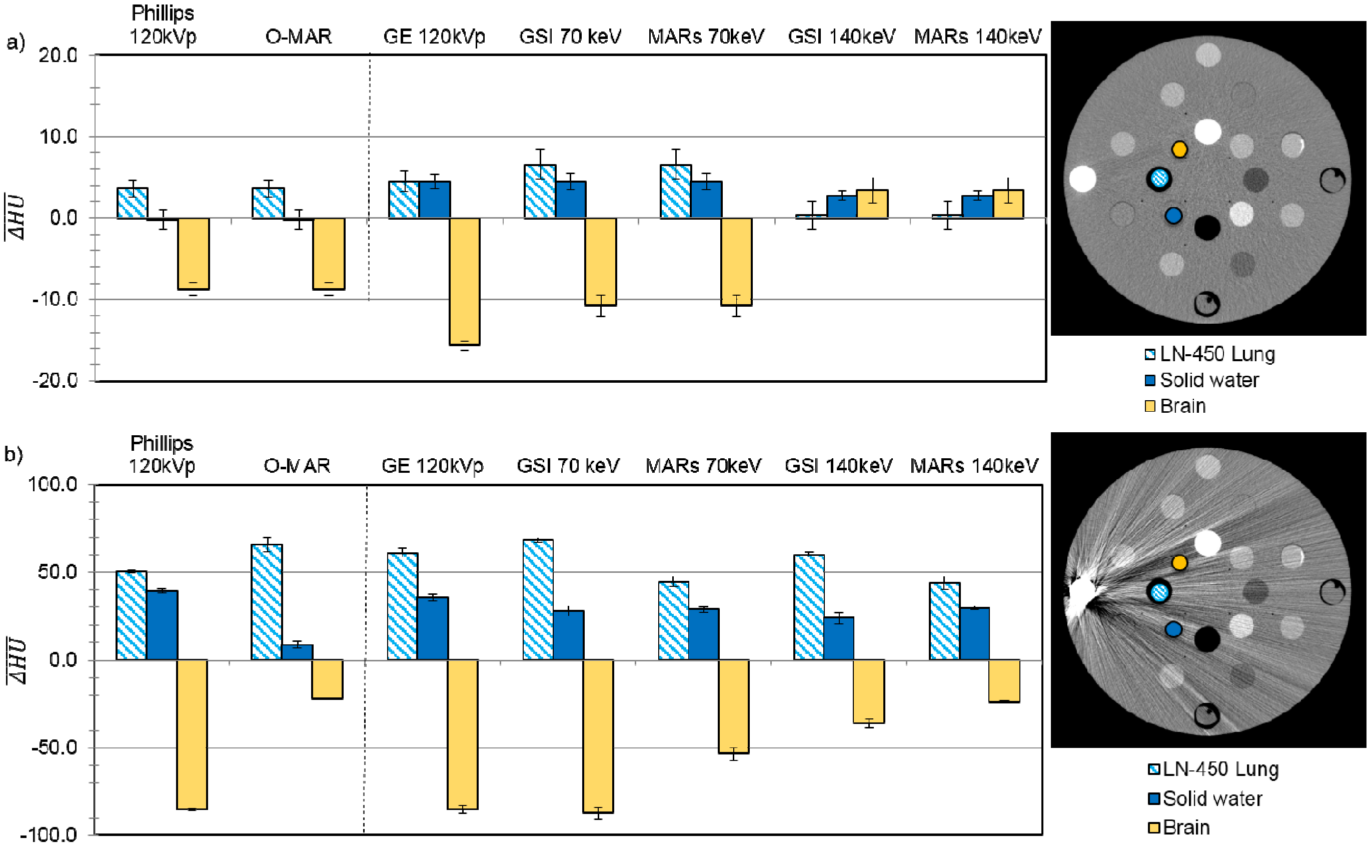


Figure B.1: HU errors ( ) for select tissue substitute regions of interest in the RMI phantom scanned with a unilateral (a) aluminum plug and (b) cerrobend plug. are grouped by imaging technique, including uncorrected imaging methods (120kVp) and metal artifact reduction methods. For each plot, a CT image on the right (Philips 120kVp, WL =0, WW =500) shows the location of the tissue substitute inserts for which is plotted and the position of the metal insert in the phantom. Error bars indicate the standard error of the mean for three repeated scans.

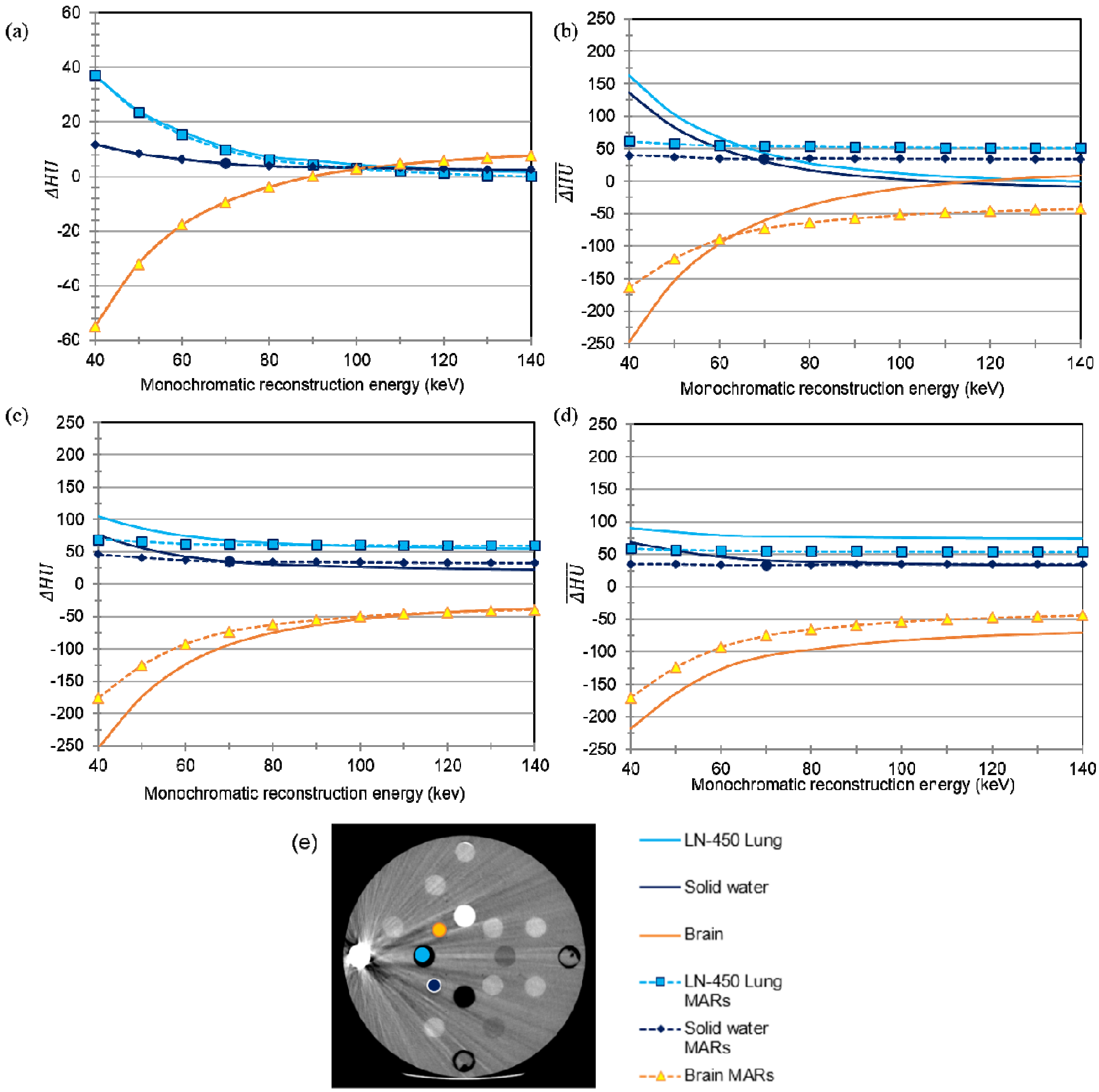


Figure B.2: HU error (   ) as a function of GSI monochromatic reconstruction energy (keV) for select tissue substitute regions of interest in the RMI phantom scanned with a unilateral (a) aluminum, (b) titanium, (c) stainless steel, and (d) cerrobend plug. Also plotted is the HU error after the GSI-MARs algorithm is applied to the virtual monochromatic images. Panel (e) shows the location of the tissue substitute inserts for which   is plotted and the position of the metal insert in the phantom.

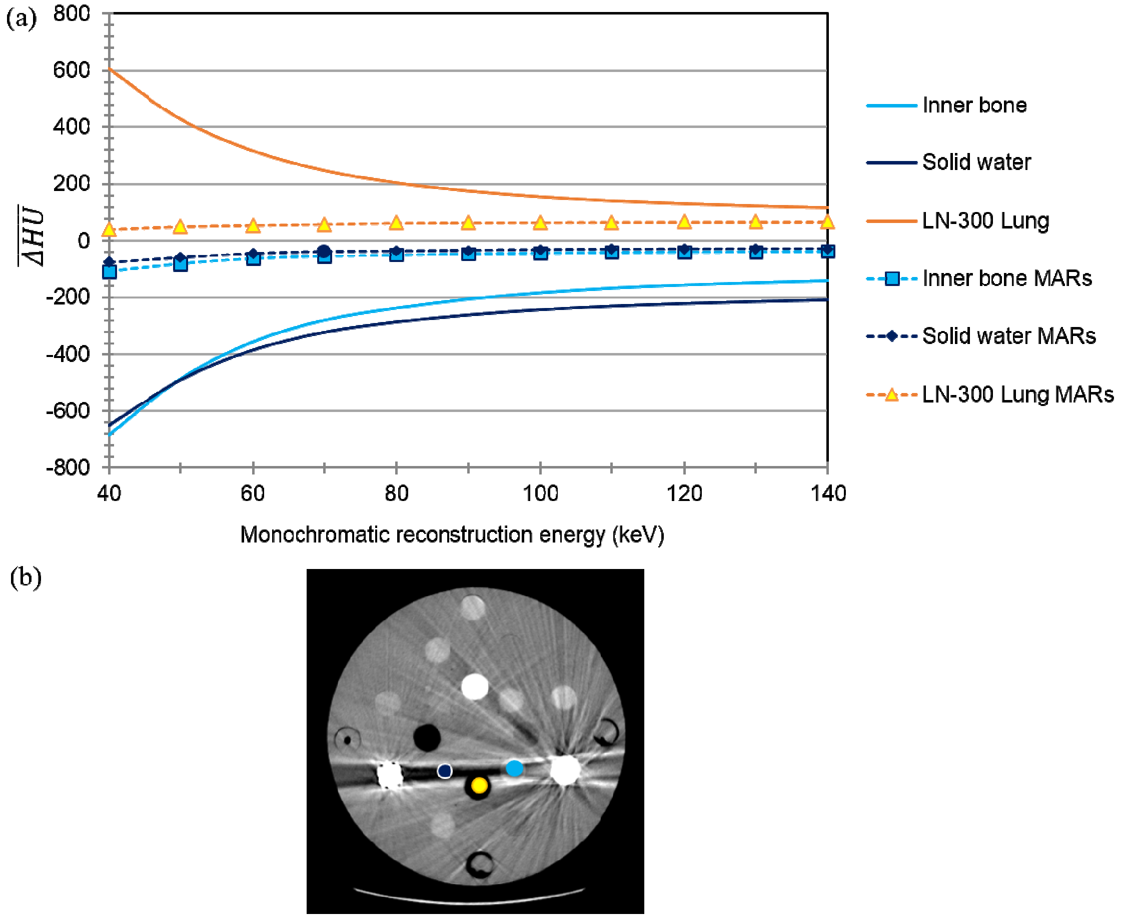


Figure B.3: HU error ( ) as a function of GSI monochromatic reconstruction energy (keV) for select tissue substitute regions of interest in the RMI phantom scanned with bilateral stainless steel and titanium inserts. Also plotted is the HU error after the GSI-MARs algorithm is applied to the monochromatic images. Panel (b) shows the location of the tissue substitute inserts for which is plotted and the position of the metal inserts in the phantom.

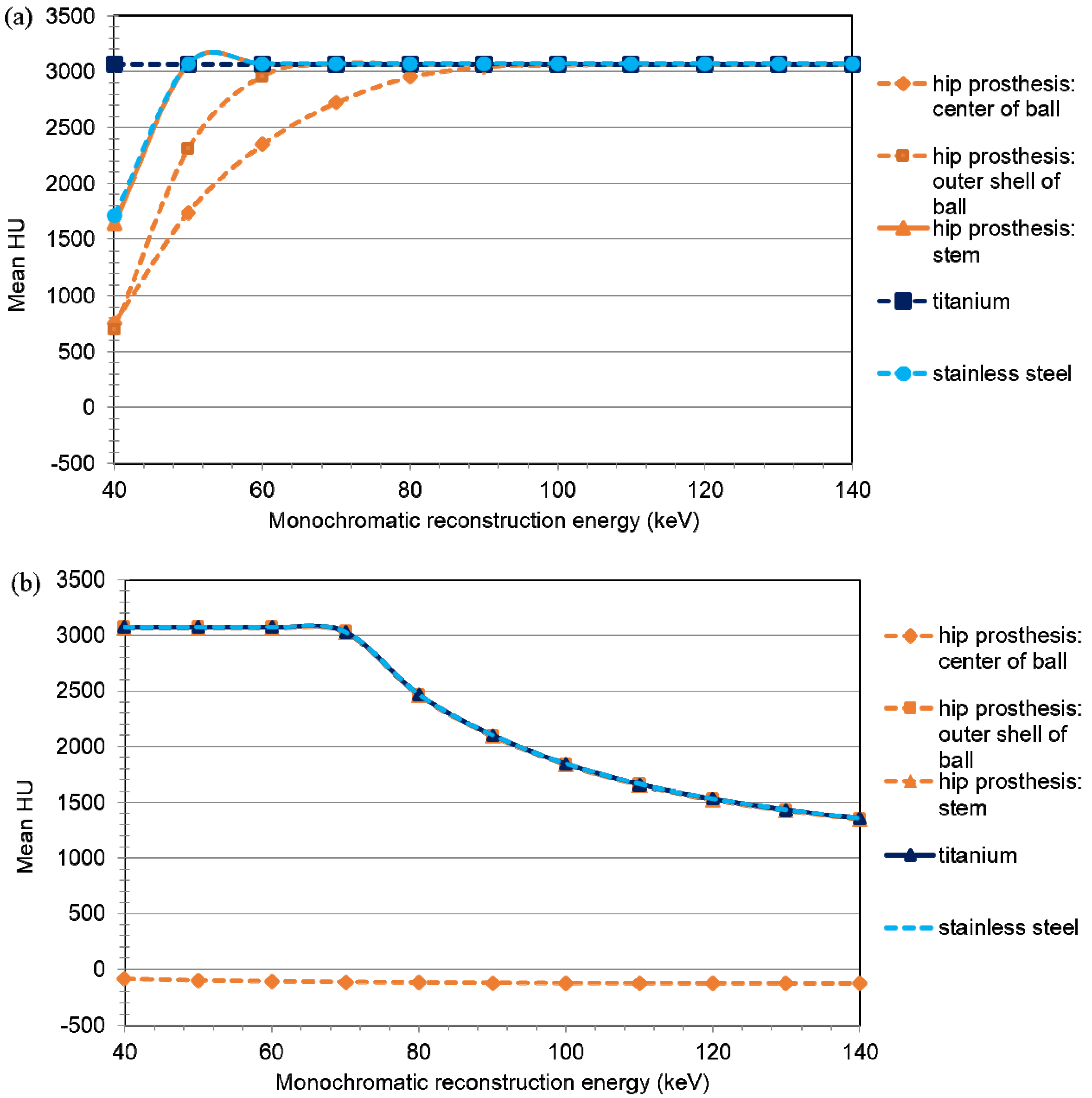


Figure B.4: Mean CT number (HU) as a function of GSI monochromatic reconstruction energy (keV) for select metal regions of interest for (a) GSI imaging without MARs and (b) GSI imaging with MARs. Metal ROIs include the center of the femoral head, the outer portion of the femoral head, and the stem of a Co-Cr hip prosthesis scanned with the anthropomorphic pelvic phantom, as well as titanium and steel inserts scanned with the Gammex RMI phantom.

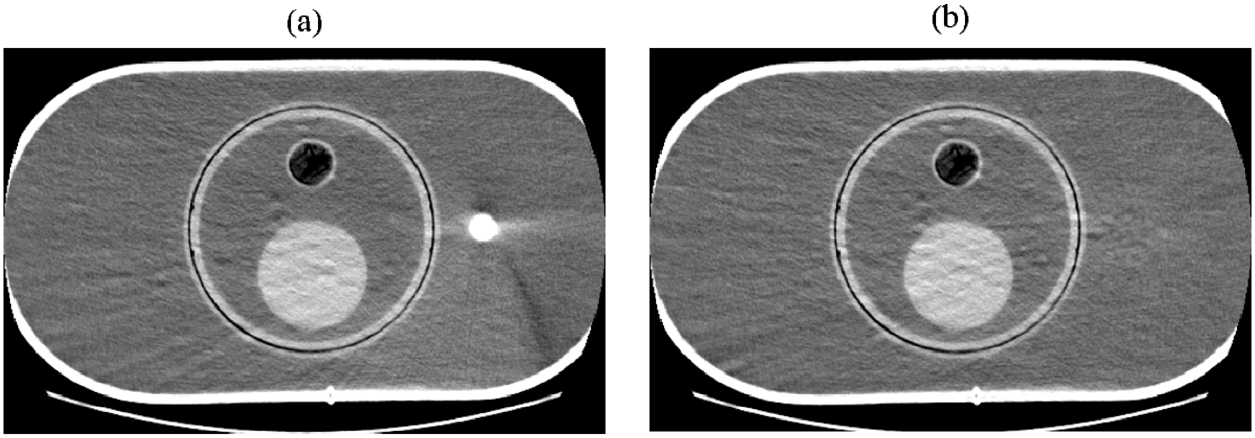


Figure B.5: CT image of the most superior aspect of the femoral head portion of a hip prosthesis for (a) GSI imaging without MARs and (b) GSI imaging with MARs applied.

Appendix C: Mobius3D beam model commissioning data

In order to perform the Mobius 3D v1.3.1 (Mobius Medical Systems, Houston, TX) dose calculations described in chapters 5 and 6, a commissioning process first had to be performed. The Mobius3D system has several stock beam models for existing models of linear accelerators. The stock beam model for a Varian Clinac 2100 was customized by inputting dosimetry parameters into Mobius3D's automodeling feature. The resulting customized beam model was then validated with various verification treatment plans. For a simple water phantom, dose calculations performed with the customized beam model were compared against measured data, and the agreement was generally very good (Table C.1). Additionally, verification plans for various anthropomorphic phantoms and patient image sets, ranging in complexity from a single photon beam to IMRT, were calculated using the customized beam model and compared to dose calculations performed with our institution's clinical Pinnacle treatment planning system. The agreement for these verification plans was very good, with all of the plans resulting in ≥ 96% of pixels passing 3D gamma analysis at ±3%/3mm criteria.

Table C.1: The % difference between Mobius3D-calculated data (using the customized beam model for our Varian Clinac 2100 linear accelerator) and measured data.

Depth [cm]	PDD values: % difference			
	5x5	10x10	15x15	20x20
5	-0.5%	-0.6%	-0.8%	-0.5%
10	-1.8%	-0.5%	-0.3%	-0.2%
15	-1.7%	0.0%	0.2%	1.3%
20	-1.7%	-0.7%	1.3%	1.9%
25	-2.0%	-0.4%	1.2%	3.6%

Output factors: % difference				
4x4	5x5	10x10	15x15	20x20
1.2%	1.1%	0.0%	-0.4%	-0.3%

Off-axis factors: % difference

Distance from CAX [cm]	Depth [cm]				
	1.5	5.0	10.0	15.0	20.0
1.0	0.2%	0.2%	0.2%	0.1%	0.1%
2.0	0.1%	0.0%	0.1%	0.1%	0.1%
3.0	0.1%	0.1%	0.1%	0.0%	0.1%
4.0	0.1%	0.1%	0.1%	0.1%	0.1%
5.0	-0.3%	-0.1%	0.0%	0.0%	0.2%
6.0	-0.5%	-0.1%	0.0%	0.0%	0.3%
8.0	-0.3%	0.0%	0.2%	0.3%	0.4%
10.0	-0.3%	0.1%	0.2%	0.7%	0.6%
12.0	-0.5%	0.2%	0.3%	0.5%	0.8%
14.0	-0.6%	0.2%	0.4%	0.7%	0.4%
16.0	-0.7%	0.0%	0.2%	0.2%	0.3%
18.0	-0.5%	0.2%	-0.2%	-0.1%	0.0%
20.0	-6.0%	0.3%	-0.3%	-0.3%	-0.3%

Table C.2 Comparison between Mobius3D and Pinnacle dose calculations for various verification plans. The % difference (Mobius3D – Pinnacle) for the mean dose to the target structure is reported, as well as the results of 3D gamma analysis.

		Target coverage		3D gamma analysis (3%/3mm)
Plan name	Treatment technique	Structure	Mean dose % difference	% pixels passing
Prostate #1	3DCRT	PTV	-1.5	99.0
Thoracic #1	3DCRT	PTV	1.4	99.9
Prostate #2	IMRT	PTV	-2.4	99.6
Thoracic #2	IMRT	PTV	2.6	99.3
Breast	3DCRT (FIF)			97.1
CNS #1	IMRT	PTV 60 Gy	1.2	97.6
Pedi Abdomen	IMRT	PTV	1.8	98.7
CNS #2	IMRT	PTV	0.2	99.4
Thoracic #3	IMRT	PTV	0.3	98.9
HN	IMRT	CTV 72 Gy	-0.9	96.7

Appendix D: Supplementary film analysis data for clinical cases

In chapter 6, the impact of metal artifact reduction methods and the use of metal kernels in the dose calculation algorithm were evaluated using two clinical cases using anthropomorphic phantoms and clinical treatments plans.

For the case of spinal hardware, the stereotactic spine radiosurgery (SSRS) treatment plan was delivered twice to an anthropomorphic thoracic phantom with titanium spinal rods. The irradiated films were then analyzed by taking dose profiles (Figure D.1) and performing 2D gamma analysis (Figures D.2-D.7). In agreement with our ion chamber results discussed in chapter 6, O-MAR and GSI imaging had little effect on the agreement between calculated and measured dose, while MARs substantially worsened the agreement. For one of our analyzed films, MARs decreased the passing rate from 79.4% to 41.5% for ±2%/2mm criteria (Figure D.7), in comparison to uncorrected 120kVp imaging. Use of metal kernels generally improved the agreement but the effect was typically small (increased the passing rate by approximately 5%).

For the case of dental fillings, a base of tongue treatment plan was created and delivered twice to the phantom with amalgam fillings in place. As with the spine case, dose profiles were analyzed (Figure D.8) and 2D gamma analysis was performed to compare measured dose against calculated dose (Figures D.9-D.14). Though the changes in passing rate due to CT metal artifact reduction were generally small, we do see some improved agreement in the gamma index maps due to application of the O-MAR and MARs algorithms. For MARs especially, there is improved agreement in the anterior portion of the analysis region of interest, though MARs also caused worsened agreement in the posterior portion of the oral cavity (Figures D.13-D.14). Additionally, to focus on the oral cavity region, which is an organ at risk in these treatments, some comparisons were also performed using a region of interest that only includes the oral cavity and disregards the dose to the teeth (Figure D.15). This smaller analysis region highlights that although MARs reduced the disagreement caused by streak artifacts, it did decrease accuracy in the posterior portion of the oral cavity, in agreement with ion chamber results (chamber 4). With this smaller analysis region, O-MAR was found to slightly improve accuracy.

Spine case

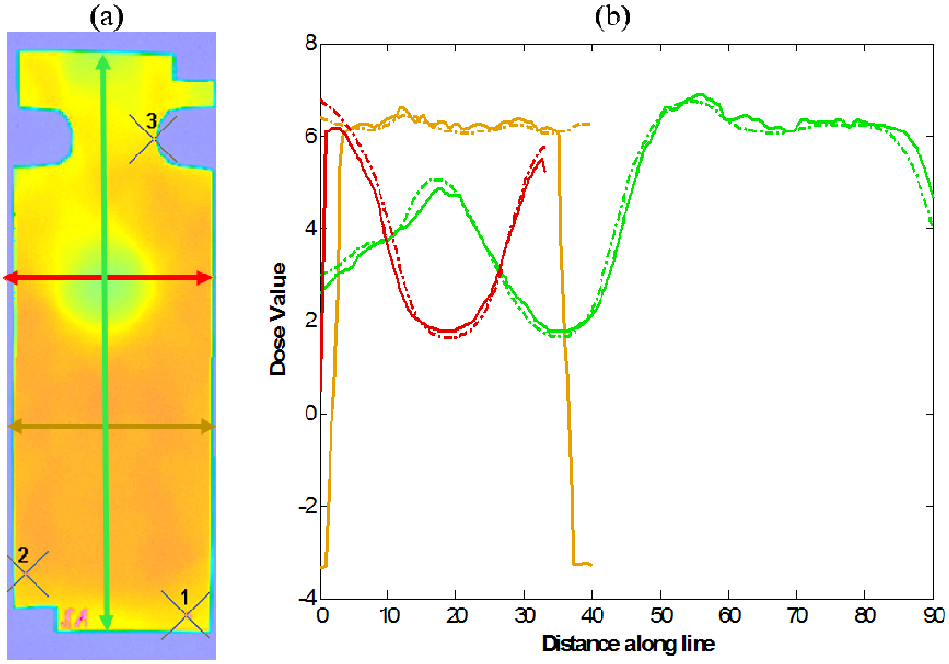
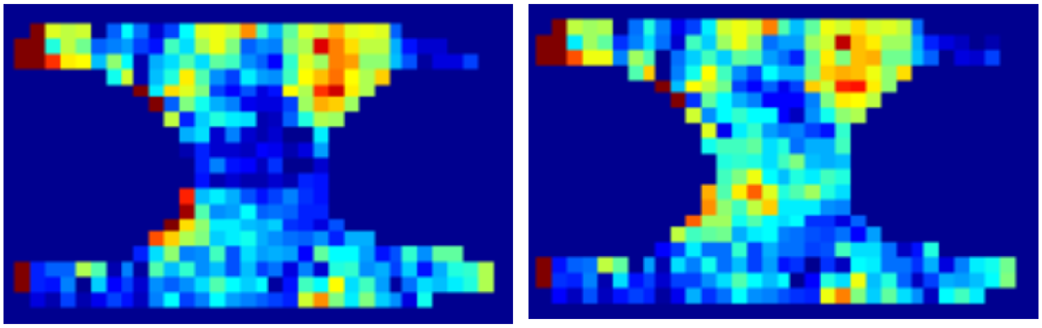
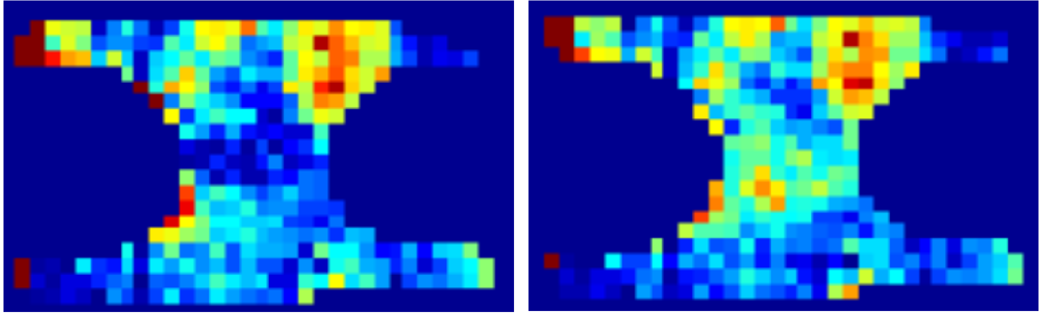


Figure D.1: (a) Digital image of radiochromic film irradiated in the spine phantom with the titanium rods in place and (b) profiles at several locations illustrating the agreement between the film-measured dose (solid line) and the Mobius3D-calculated dose (dotted line) for uncorrected baseline CT images and no metal kernels.



(a) Baseline: 77.2%

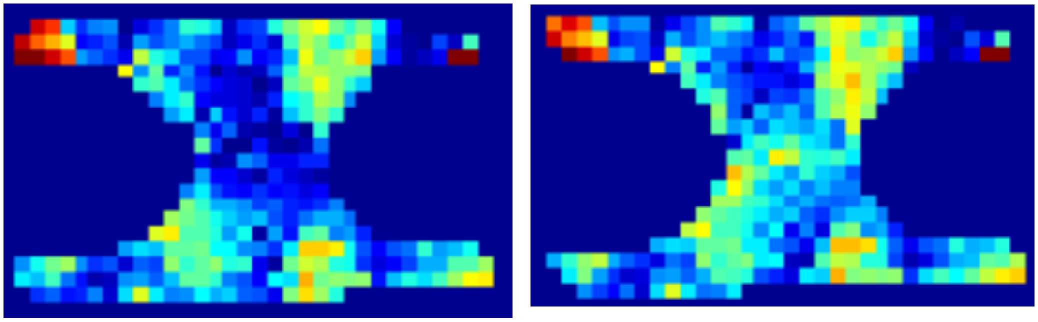
(b) O-MAR : 76.5%



(c) Baseline + MK: 76.6%

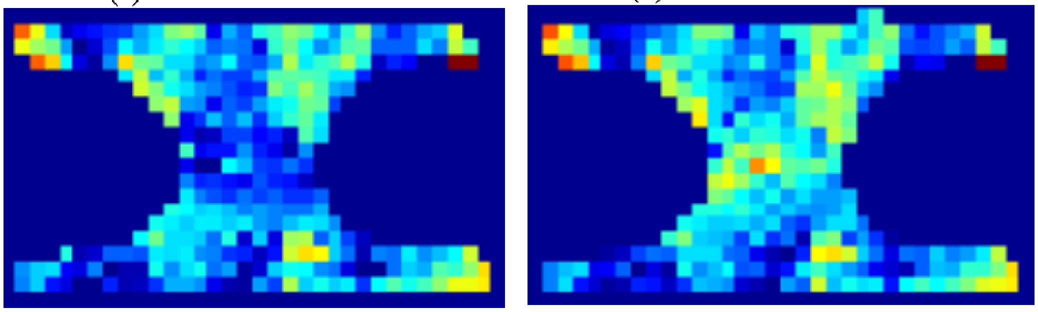
(d) O-MAR + MK: 75.2%

Figure D.2: Gamma analysis results (±2%/2mm criteria) for film#1. Mobius 3D dose calculations were performed using baseline (uncorrected Philips 120kVp) images as well as O-MAR images, with and without metal kernels (MK).



(a) Baseline: 84.9%

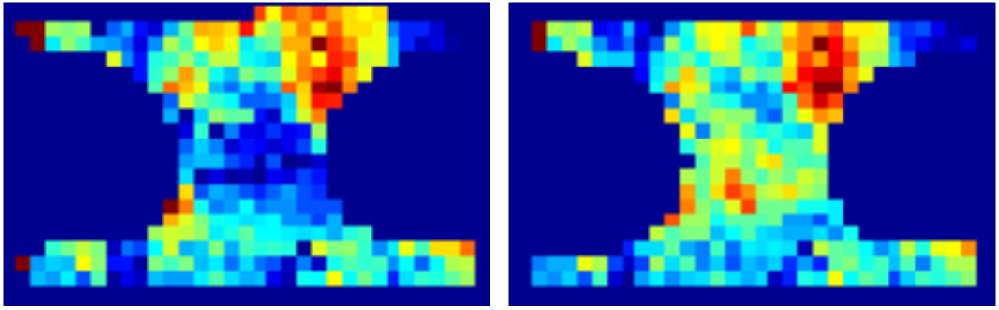
(b) O-MAR : 79.4%



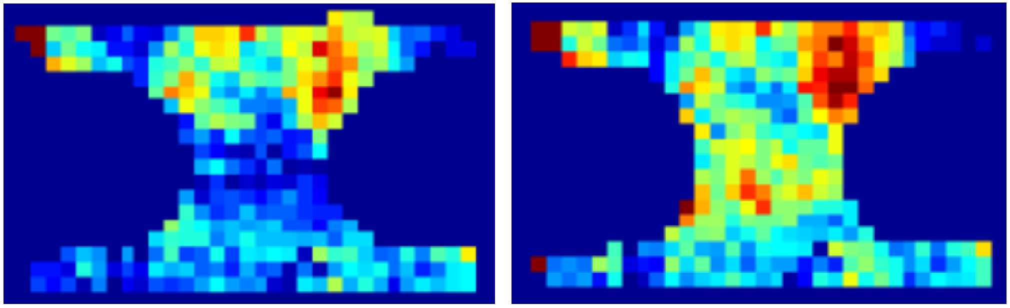
(c) Baseline + MK: 89.2%

(d) O-MAR + MK: 85.2%

Figure D.3: Gamma analysis results (±2%/2mm criteria) for film#2. Mobius 3D dose calculations were performed using baseline (uncorrected Philips 120kVp) images as well as O-MAR images, with and without metal kernels (MK).

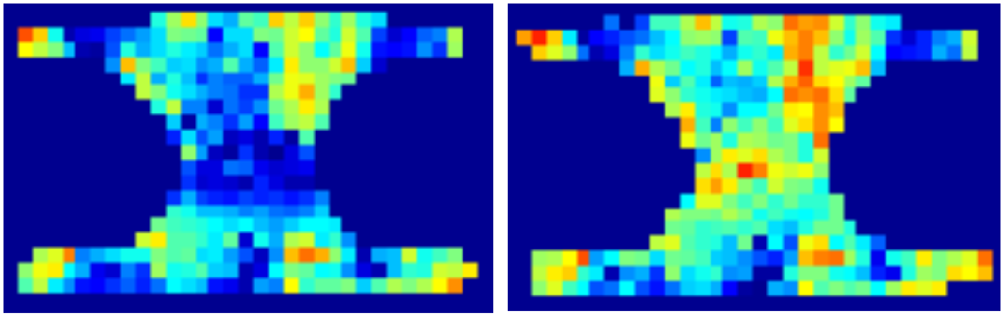


(a) Baseline: 69.2% (b) GSI : 63.1%

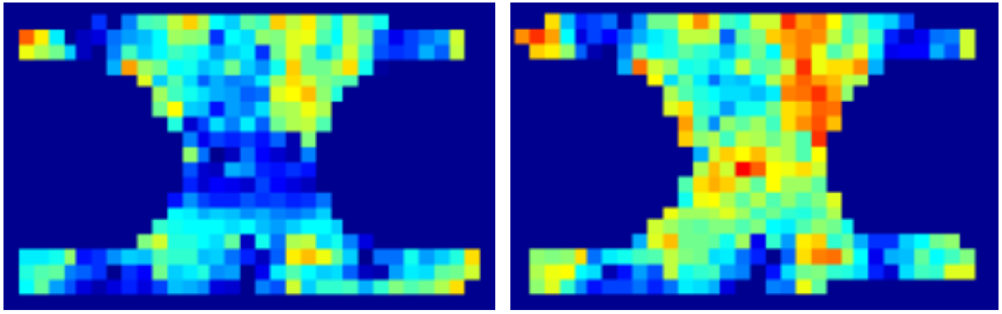


(c) Baseline + MK: 78.2% (d) GSI + MK: 62.1%

Figure D.4: Gamma analysis results (±2%/2mm criteria) for film#1. Mobius 3D dose calculations were performed using baseline (uncorrected GE 120kVp) images as well as GSI 140keV images, with and without metal kernels (MK).

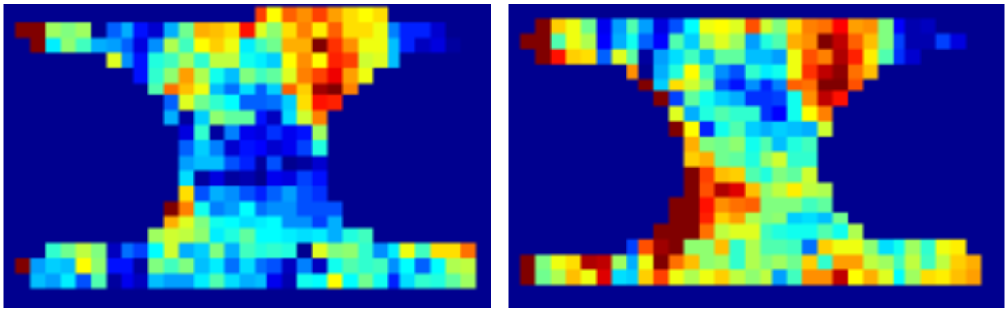


(a) Baseline: 79.4% (b) GSI : 61.8%



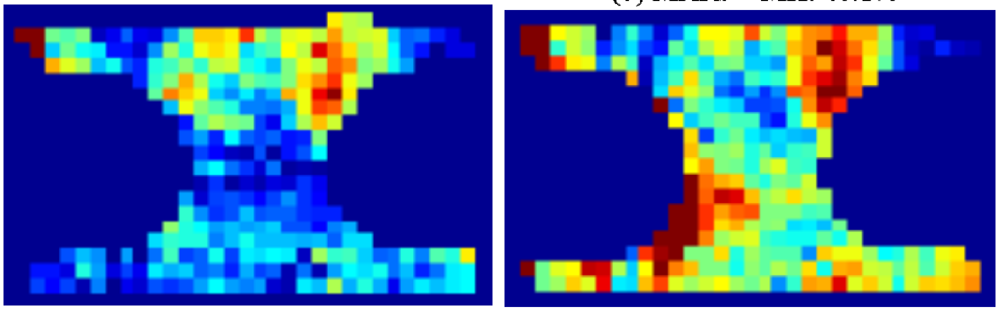
(c) Baseline + MK: 83.3% (d) GSI + MK: 58.4%

Figure D.5: Gamma analysis results (±2%/2mm criteria) for film#2. Mobius 3D dose calculations were performed using baseline (uncorrected GE 120kVp) images as well as GSI 140keV images, with and without metal kernels (MK).



(a) Baseline: 69.2%

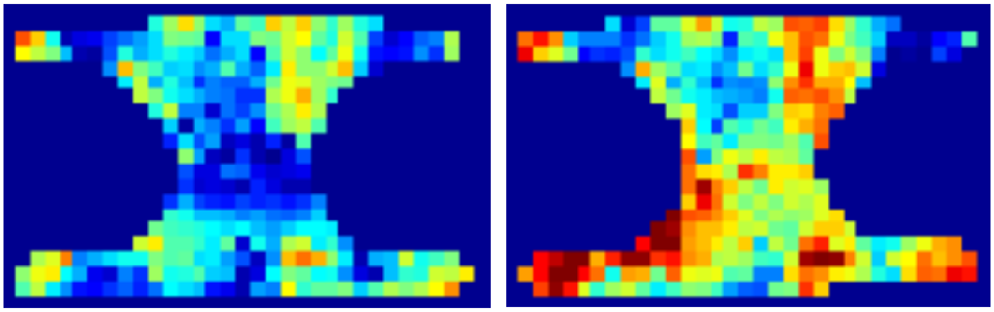
(b) MARs : 45.6%



(c) Baseline + MK: 78.2%

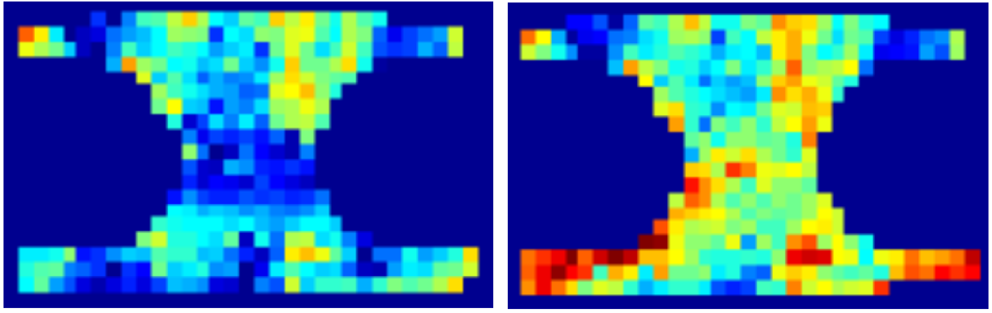
(d) MARs + MK: 46.1%

Figure D.6: Gamma analysis results (±2%/2mm criteria) for film#1. Mobius 3D dose calculations were performed using baseline (uncorrected GE 120kVp) images as well as MARs 140keV images, with and without metal kernels (MK).



(a) Baseline: 79.4%

(b) MARs : 41.5%



(c) Baseline + MK: 83.3%

(d) MARs + MK: 46.6%

Figure D.7: Gamma analysis results (±2%/2mm criteria) for film#2. Mobius 3D dose calculations were performed using baseline (uncorrected GE 120kVp) images as well as MARs 140keV images, with and without metal kernels (MK).

Dental case

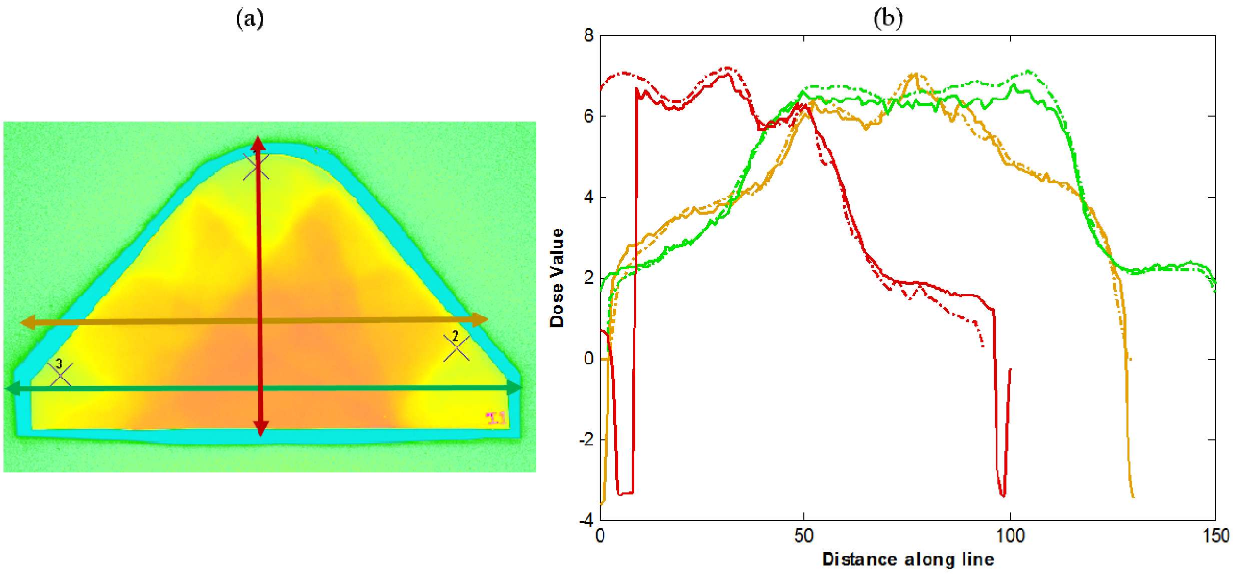
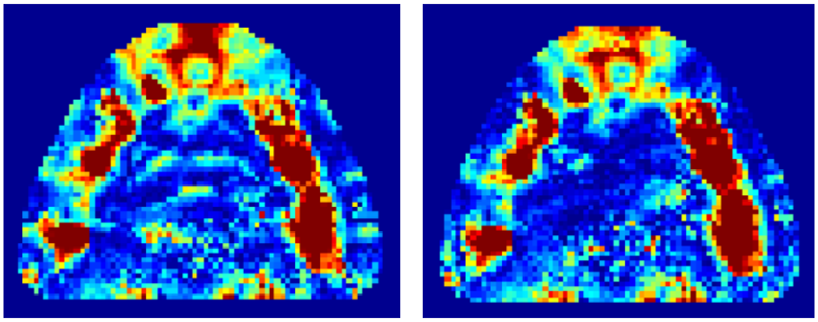
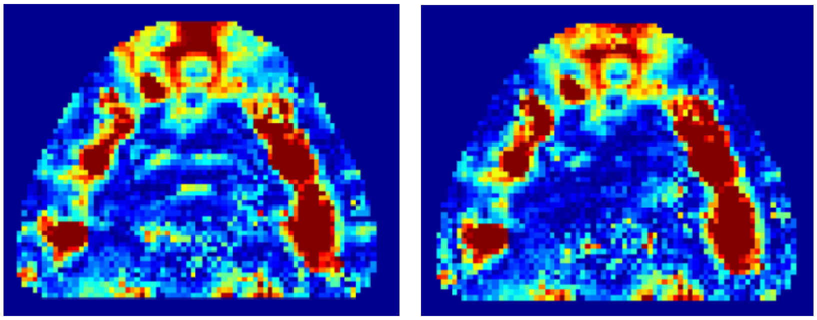


Figure D.8: Digital image of radiochromic film irradiated in the dental phantom with amalgam fillings in place and (b) profiles at several locations illustrating the agreement between the film-measured dose (solid line) and the Mobius3D-calculated dose (dotted line) for uncorrected baseline CT images and no metal kernels.



(a) Baseline: 70.2%

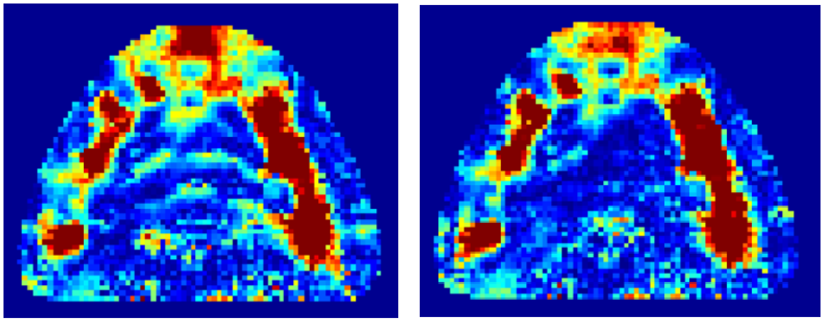
(b) O-MAR: 70.9%



(c) Baseline + MK: 70.1%

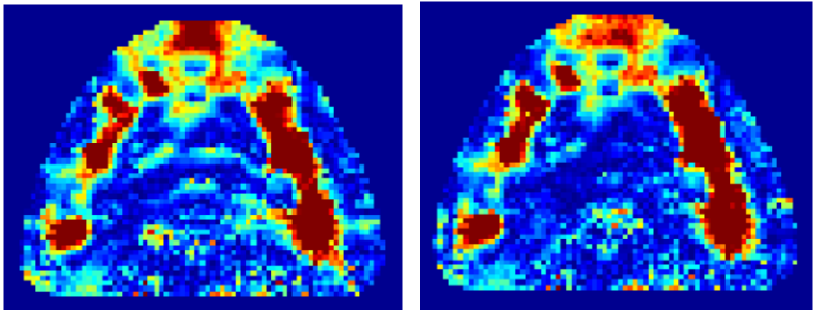
(d) O-MAR + MK: 70.8%

Figure D.9: Gamma analysis results (±2%/2mm criteria) for film#1. Mobius 3D dose calculations were performed using baseline (uncorrected Philips 120kVp) images as well as O-MAR images, with and without metal kernels (MK).



(a) Baseline: 69.5%

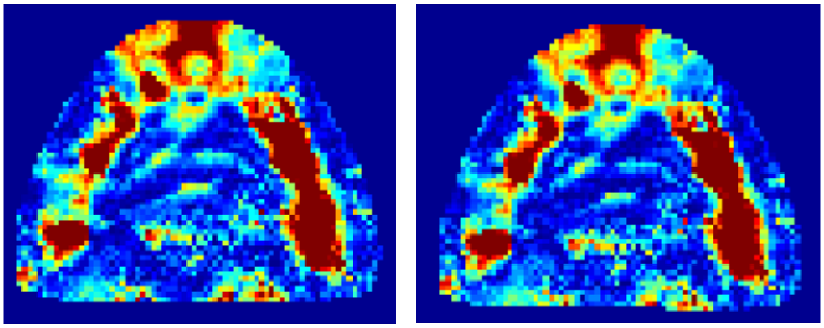
(b) O-MAR: 73.1%



(c) Baseline + MK: 70.8%

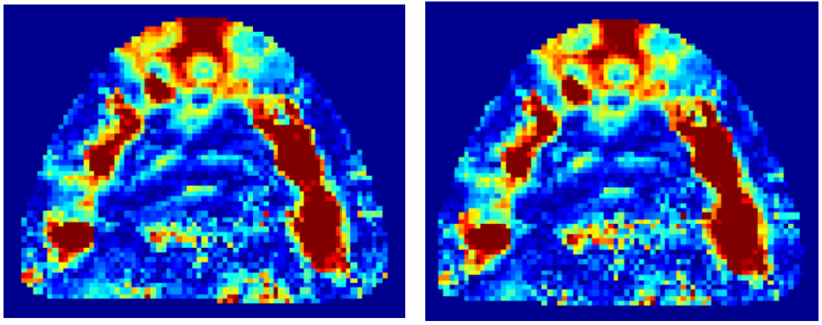
(d) O-MAR + MK: 72.8%

Figure D.10: Gamma analysis results (±2%/2mm criteria) for film#2. Mobius 3D dose calculations were performed using baseline (uncorrected Philips 120kVp) images as well as O-MAR images, with and without metal kernels (MK).



(a) Baseline: 67.6%

(b) GSI: 69.2%

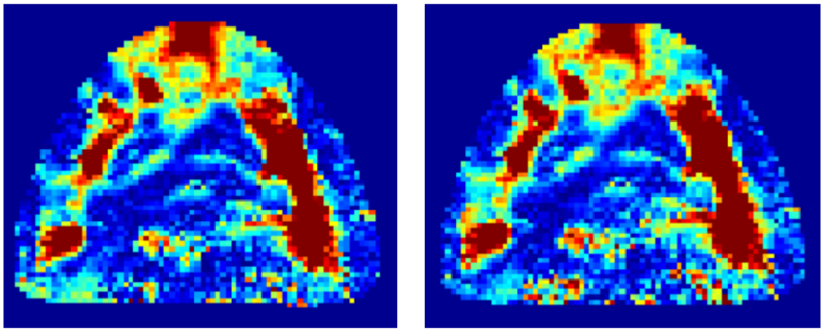


(c) Baseline + MK: 69.3%

(d) GSI + MK: 68.4%

Figure D.11: Gamma analysis results (±2%/2mm criteria) for film#1. Mobius 3D dose calculations were performed using baseline (uncorrected Philips 120kVp) images as well as GSI 140keV images, with and without metal kernels (MK).

(a) Baseline: 68.6% (b) GSI: 65.5%



(c) Baseline + MK: 68.7% (d) GSI + MK: 64.9%

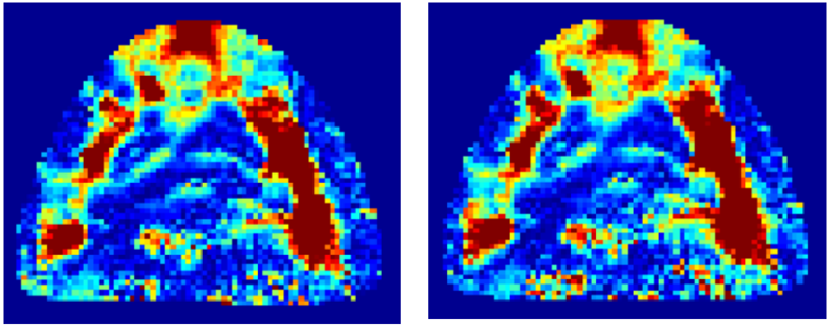
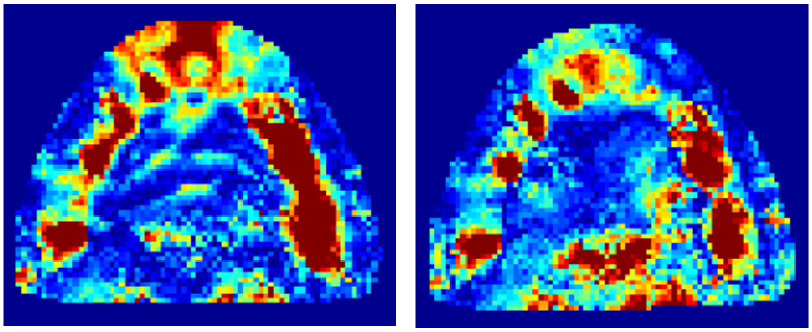


Figure D.12: Gamma analysis results (±2%/2mm criteria) for film#2. Mobius 3D dose calculations were performed using baseline (uncorrected Philips 120kVp) images as well as GSI 140keV images, with and without metal kernels (MK).

(a) Baseline: 67.6% (b) MARs: 68.5%



(c) Baseline + MK: 69.3% (d) MARs + MK: 67.4%

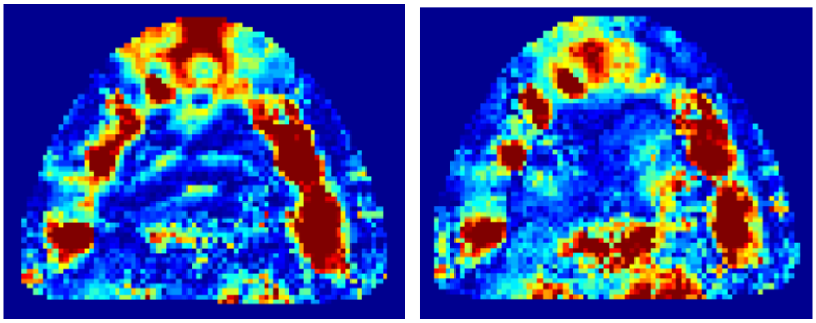
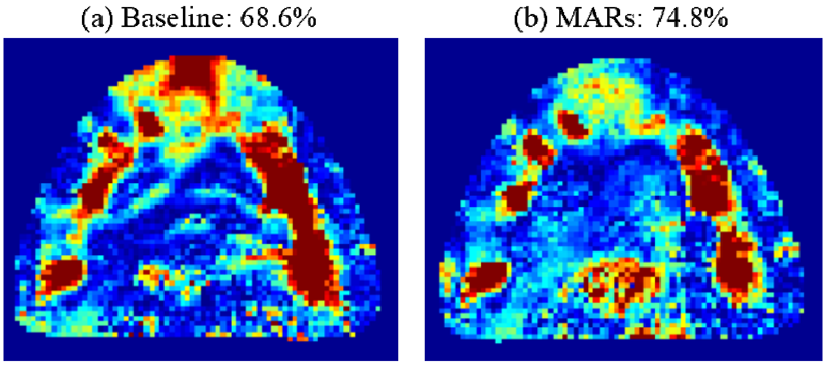
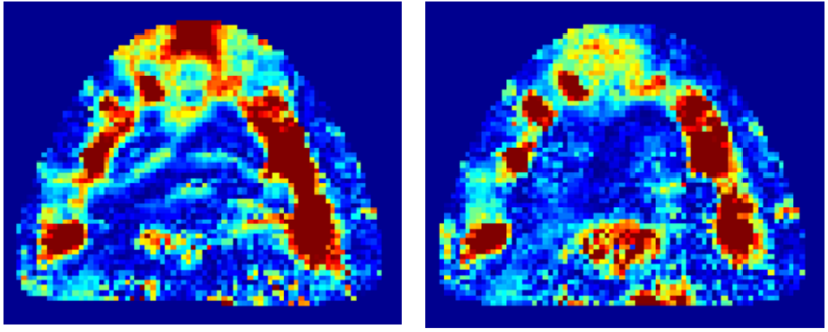


Figure D.13: Gamma analysis results (±2%/2mm criteria) for film#1. Mobius 3D dose calculations were performed using baseline (uncorrected Philips 120kVp) images as well as MARs 140keV images, with and without metal kernels (MK).

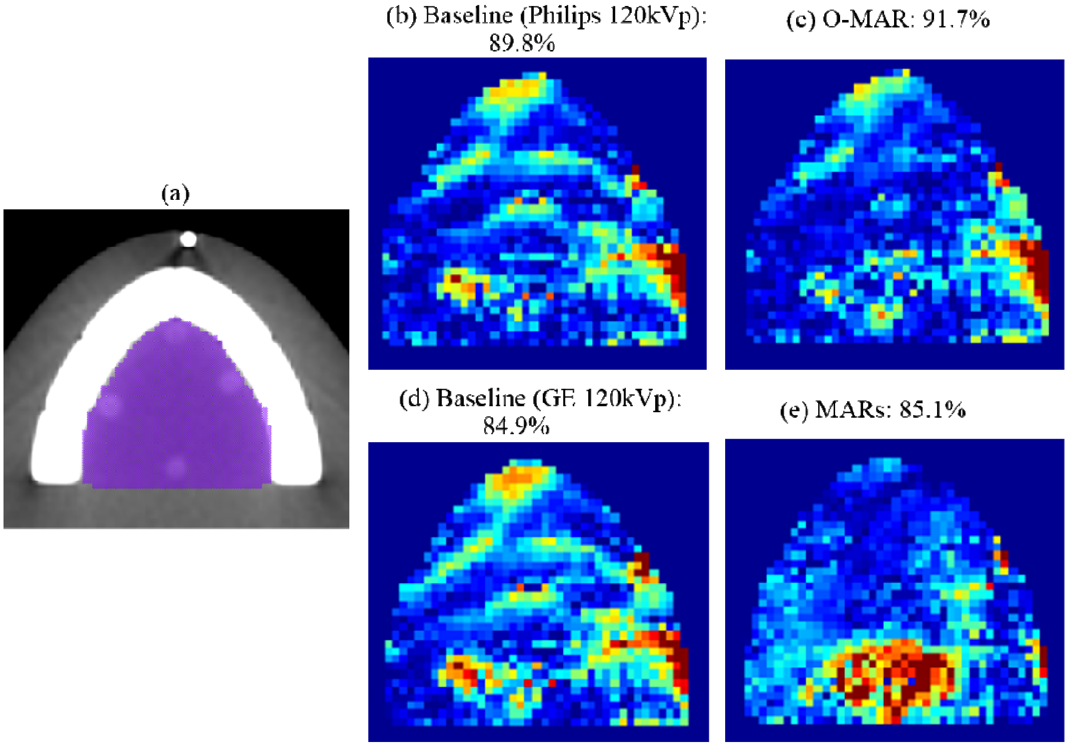

(a) Baseline: 68.6%

(b) MARs: 74.8%


(c) Baseline + MK: 68.7%

(d) MARs + MK: 73.5%

Figure D.14: Gamma analysis results (±2%/2mm criteria) for film#2. Mobius 3D dose calculations were performed using baseline (uncorrected Philips 120kVp) images as well as MARs 140keV images, with and without metal kernels (MK).


(a)

(b) Baseline (Philips 120kVp): 89.8%

(c) O-MAR: 91.7%

(d) Baseline (GE 120kVp): 84.9%

(e) MARs: 85.1%

Figure D.15: Film gamma analysis results for the dental phantom irradiated with amalgam fillings using an analysis region of interest (purple) that focuses on the oral cavity (Panel a). Panels (b)-(e) show results for ±2%/2mm criteria using dose calculations performed with uncorrected CT images (Philips 120kVp and GE 120kVp) and metal artifact reduction (O-MAR and MARs).

Vita

Jessie Yanshu Huang-Vredevoogd was born in Lian Chen, China on September 17, 1987 to parents Lidong Huang and Dexin Zhang. After graduating from A&M Consolidated High School in College Station, TX in May 2006, she attended Rice University in Houston, Texas. In May 2010, she graduated summa cum laude with a Bachelor of Science in Physics with a concentration in biological physics. She then entered The University of Texas Health Science Center at Houston Graduate School of Biological Sciences to pursue a Ph.D in Medical Physics. Jessie is married to Kevin Vredevoogd and enjoys cooking, traveling, and recycling.

Permanent address:
810 Holston Hills. Dr.
College Station, TX 77845